TROPOSPHERIC O_3 MODELING STUDY: CONTRIBUTIONS OF ANTHROPOGENIC AND BIOGENIC SOURCES TO O_3-CO AND O_3-CH_2O CORRELATIONS

A Dissertation
Presented to
The Academic Faculty

by

Ye Cheng

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Earth and Atmospheric Sciences

Georgia Institute of Technology
May 2018

TROPOSPHERIC O_3 MODELING STUDY: CONTRIBUTIONS OF ANTHROPOGENIC AND BIOGENIC SOURCES TO O_3-CO AND O_3-CH_2O CORRELATIONS

Approved by:

Dr. Yuhang Wang, Advisor
School of Earth and Atmospheric Sciences
Georgia Institute of Technology

Dr. Rodney Weber
School of Earth and Atmospheric Sciences
Georgia Institute of Technology

Dr. Yi Deng
School of Earth and Atmospheric Sciences
Georgia Institute of Technology

Dr. James Crawford
Langley Research Center
National Aeronautics and Space Administration

Dr. Nga Lee Ng
School of Earth and Atmospheric Sciences
Georgia Institute of Technology

Date Approved: April 2, 2018

ACKNOWLEDGEMENTS

I would like to start by thanking my advisor, Dr. Yuhang Wang, for his support and guidance throughout my Ph.D. I have been benefitted from not only his knowledge and insights, but also his passion for science and dedication to high standards. I am also grateful for my Ph.D. thesis committee members, Dr. Yi Deng, Dr. Rodney Weber, Dr. Nga Lee, and Dr. James Crawford, for their review of my dissertation and for their comments and suggestions.

Thanks also go to all the current members and alumni of Yuhang Wang's group. I have enjoyed my time working with Dr. Yongjia Song, Dr. Yuzhong Zhang, Charles Smeltzer, Yufei Zou, Tom Loadholt, Ruixiong Zhang, Hang Qu, Jianfeng Li, Ziming Ke, Aoxing Zhang and Qiyang Yan. I am also grateful for the DISCOVER-AQ science team. The data they collected are important foundations of this work. I also thank my coauthors, especially Dr. James Crawford, for the constructive and insightful discussions.

I am also grateful that I have made many good friends at Georgia Tech. I shared with them interesting discussions, fun events, and precious moments.

Finally, I would like to express my deepest gratitude to my parents and my wife for their unending love and support. This work would not have been possible without their constant encouragement throughout my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

indirect (oxidation of VOCs) from Houston ship channel emission, Houston urban emission, other kinds of anthropogenic emissions, boundary conditions, and biogenic isoprene's oxidation, respectively. In Colorado, the title of "Oil & Gas" stand for the results of emissions from oil and gas operations. In Maryland, the title "Anthro VOCs" represents the results from all anthropogenic VOCs emissions.

LIST OF SYMBOLS AND ABBREVIATIONS

SYMBOLS

CO	carbon monoxide
CH_2O	formaldehyde
O_3	ozone
NO	nitrogen monoxide
NO_2	nitrogen dioxide
NO_x	nitrogen oxides
OH	hydroxyl radical
HO_2	hydrogen peroxide
$CO_{anthroCO}$	CO from primary anthropogenic emissions
$CO_{anthroVOCs}$	CO from oxidation of anthropogenic VOCs
CO_{BC}	CO from transport from model lateral and upper boundaries
$CO_{bioISOP}$	CO from oxidation of biogenic isoprene
$CO_{background}$	CO from transport from model lateral and upper boundaries (used in O_3 estimator)
$CH_2O_{anthroVOCs}$	CH_2O from primary anthropogenic emissions and the oxidation of anthropogenic VOCs
$CH_2O_{bioISOP}$	CH_2O from oxidation of biogenic isoprene
CH_2O_{BC}	CH_2O from transport from model lateral and upper boundaries
$CH_2O_{background}$	CH_2O from transport from model lateral and upper boundaries (used in O_3 estimator)
CH_2O_{HSC}	CH_2O from oxidation of VOCs from the selected grid cells at HSC
CH_2O_{urban}	CH_2O from oxidation of VOCs from Houston urban areas
$CH_2O_{anthro\text{-}others}$	CH_2O from oxidation of VOCs from other anthropogenic sources

$O_{3\ bioISOP}$	anthropogenic O_3 production
$O_{3\ anthro}$	biogenic O_3 production
$O_{3\ bachground}$	O_3 from transported from the lateral and upper model boundaries

ABBREVIATIONS

3-D	three-dimensional
VOC	volatile organic compound
UV	ultraviolet
TIR	thermal infrared
NIR	near infrared
SCIAMACHY	SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY
MOPITT	Measurement Of Pollution In The Troposphere
OMI	Ozone Monitoring Instrument
GOME-2	Global Ozone Monitoring Experiment–2
OMPS	Ozone Mapping Profiler Suite
SOF	Solar Occultation Flux
DOAS	differential optical absorption spectroscopy
HSC	Houston ship channel
REAM	Regional chEmical trAnsport Model
WRF	Weather Research and Forecasting
CFSR	Climate Forecast System Reanalysis
NEI	National Emission Inventory
MEGAN	Model of Emissions of Gases and Aerosols from Nature

DISCOVER-AQ	Deriving Information on Surface Conditions from Column and Vertically Resolved Observations Relevant to Air Quality
LOOCV	leave-one-out cross validation
AutoGCs	Automated Gas Chromatographs
TEMPO	Tropospheric Emissions: Monitoring of Pollution
TCEQ	Texas Commission on Rnvironmental Quality

SUMMARY

Tropospheric O_3 and CO are major pollutants in the troposphere. Strong correlation between O_3 and CO was observed during the DISCOVER-AQ aircraft experiment in July 2011 over the Washington-Baltimore area. The observed correlation does not vary significantly with time or altitude in the boundary layer. The observations are simulated well by a regional chemical transport model. We analyze the model results to understand the factors contributing to the observed O_3-CO regression slope, which has been used in past studies to estimate the anthropogenic O_3 production amount. We trace separately four different CO sources: primary anthropogenic emissions, oxidation of anthropogenic VOCs, oxidation of biogenic isoprene, and transport from the lateral and upper model boundaries. Modeling analysis suggests that the contribution from biogenic isoprene oxidation to the observed O_3-CO regression slope is as large as that from primary anthropogenic CO emissions. As a result of decrease of anthropogenic primary CO emissions during the past decades, biogenic CO from oxidation of isoprene is increasingly important. Consequently, observed and simulated O_3-CO regression slopes can no longer be used directly with an anthropogenic CO emission inventory to quantify anthropogenic O_3 production over the United States. The consistent enhancement of O_3 relative to CO observed in the boundary layer, as indicated by the O_3-CO regression slope, provides a useful constraint on model photochemistry and emissions.

As an extension, we analyze the scenario of O_3-CO regression slopes in the entire United States and China regions. The O_3-CO regression slope ~ 0.3 is simulated over the eastern outflow regions over the ocean. Over the eastern inland regions of both countries,

the O_3-CO regression slope is lower than that over the outflow region, reflecting in part continuous O_3 production in the outflow region. The simulation result shows that the proportion of contribution from biogenic isoprene to the regressed O_3-CO slopes various depending on the corresponding local emission scenario. While biogenic isoprene oxidation makes a comparable contribution as anthropogenic emissions in the eastern US, the latter dominates over eastern China. Over the western inland regions of both countries, the O_3-CO regression slope can be higher than the eastern inland regions due to transport from lateral and upper boundaries. The observations of O_3-CO regression slope provide the means to understand the relative importance of anthropogenic and biogenic emissions on O_3 as well as transport.

In addition to O_3-CO, strong correlations and consistent linear regression slopes of O_3-CH_2O and CO-CH_2O were also observed during the DISCOVER-AQ aircraft experiment in July 2011 over the Washington-Baltimore area. Same as CO, we also analyze the model results to understand the factors contributing to the observed O_3-CH_2O regression slope by tracing separately three different CH_2O sources: primary secondary anthropogenic sources, biogenic isoprene oxidation, and transport from model boundaries. Results show biogenic isoprene oxidation makes the largest contribution to the regression slope of O_3-CH_2O across much of the eastern United States, providing a good indicator for O_3 enhanced by biogenic VOCs. In contrast, the regression slope of O_3-CO is controlled by both anthropogenic and biogenic emissions. Therefore, the CO-CH_2O linear relationship can be applied to track the contributions to surface O_3 by anthropogenic and biogenic factors. Making use of these linear dependences, we build a fast-response ozone estimator using near surface CH_2O and CO concentrations as inputs. We examine the quality of this

O_3 estimator by increasing or decreasing anthropogenic emissions by up to 50%. The estimated O_3 distribution is in reasonably good agreement with the full-model simulations (R^2 >0.77 in the range of -30% to +50% of anthropogenic emissions). The analysis provides the basis for using high-quality geostationary satellites with UV, thermal infrared, or near infrared instruments for observing CH_2O and CO to improve surface O_3 distribution monitoring. The estimation model also provides 6 observation-derived regional metrics to evaluate and improve full-fledged 3-D air quality models.

The NASA DISCOVER-AQ airborne campaigns were also carried out around the Houston and Denver metropolitan areas in the summers of 2013 and 2014, respectively. Using the 2011 national emissions inventory (NEI), a regional chemical transport model (REAM) is applied to analyze the aircraft observations. We find that major model discrepancies are driven by large underestimates of alkane emissions in both regions. Modeling analysis suggests increases of alkane emissions by a factor of 15 in the Houston Ship Channel, where ship-transport, ship-unloading, storage, domestic transportation of oil take place, and by a factor of 5 in the regions of oil and gas exploration of Denver. The large increase of alkane emissions has drastically different effects on O_3 concentrations depending on the strength of biogenic emissions. A useful metric to diagnose the effects of alkane emissions on photochemistry is the least-squares regression slope of O_3 to CH_2O, which increases by 30% and 80% in Houston and Denver, respectively, due to the increases of alkane emissions, leading to good agreement between model simulations and aircraft observations. Our finding implies that alkane emissions from oil and gas related sources may be substantially underestimated by the NEI, leading to corresponding underestimates of anthropogenic contributions to O_3 particularly over the western United States where biogenic VOC emissions are low. In regions like Denver, reducing alkane emissions is urgently required to control summertime O_3 concentrations.

CHAPTER 1. INTRODUCTION

1.1 Statement of the Problem

Correlations between chemical species can be used effectively to diagnose atmospheric processes. These correlations can be affected by many atmospheric chemical and physical processes, for example, chemical reactions, surface emissions, mixture of air (e.g. transport), and metrological conditions (e.g. boundary layer height). Study on the correlation between ozone (O_3) and other species is challenging because O_3 is secondary produced and normally has non-linear correlations with its various precursors. A good understanding of the correlation of O_3 and other species can provide an important piece of information for regional air pollution control.

In this dissertation, I will focus on three topics pertaining to the study of correlations of O_3 and other species: (1) The contributions of various sources to the correlations of O_3 with other species, (2) using the correlations to estimate lower tropospheric O_3 concentrations and (3) the regional anthropogenic O_3 precursor emissions and its effect on the correlations. In the first topic, I will investigate the contributions of the anthropogenic and biogenic sources to the correlations of O_3 with carbon monoxide (CO) and formaldehyde (CH_2O). In the second topic, I will develop a method of surface O_3 estimation with the concentrations of CO and CH_2O. In the last topic, I will focus on the volatile organic compounds (VOCs) emissions correction at Houston areas and study its effect on reginal O_3- CH_2O correlation.

1.1.1 The Contributions of Different Sources to the Correlations of O_3 with other Species

Tropospheric O_3 and CO burdens depend on a wide range of physical and chemical processes that take place in the troposphere [Logan et al., 1981; Lelieveld and Dentener, 2000]. Both of these two species exert influence on the tropospheric oxidizing capacity and affecting the air quality and climate. Their relatively long lifetimes enable these two species to undergo intercontinental transport which exerts upon the chemistry of the troposphere, especially with regard to the ozone budget. This effect is particularly important in the North Atlantic because during the summer the Azores-Bermuda high-pressure system tends to circulate the exported pollution from Northeastern United States which is one of the world's largest emission regions. Thus, the emission characters from source regions can significantly affect atmospheric composition in downwind regions. Li et al. [2002] found that North United States anthropogenic emissions make contribution to 20% of summer violations of O_3 air quality standards in Europe.

In the past, an evidential enhancement ratio (slope) of O_3 to CO of about 0.3 have been found from both observations and model simulation [Parrish et al., 1993; Chin et al., 1994; Parrish et al., 1998; C´ardenas et al., 1998; Li et al., 2002; Honrath et al., 2004; Mao and Talbot, 2004]. Aircraft measurements, employing northeastern United States and the pollutant outflow region in the northern Atlantic as the most prevailed studied areas, show that O_3-CO correlations in the summer are positive and often strong, depending on the weather patterns during the measurement period [Buhr et al., 1996; Daum et al., 1996; Cooper et al., 2002a, b; Huntrieser et al., 2005]. Researchers above attribute this intimate correlation to the synchronous anthropogenic emissions of CO and nitrogen oxides (NO_x) which is one of the most important precursors of O_3. Since CO has long been used as one

of typical anthropogenic tracers [Duncan et al., 2007; Holloway et al., 2000; Khalil and Rasmussen, 1994; Seinfeld and Pandis, 2016], the O_3-CO correlation reveals a potential application to use the slope of O_3 to CO to estimate the anthropogenic influence on O_3, though limited by the large uncertainties, for example, the O_3 titration in the presence of NO_x [Chin et al., 1994; Parrish et al., 1998; C´ardenas et al., 1998; Li et al., 2002; Mao and Talbot, 2004; Real et al., 2008] and stratospheric O_3 intrusion [Parrish et al., 1998]. However, the reasonability of this plausible application of using the O_3-CO correlation with the anthropogenic CO emissions to estimate anthropogenic O_3 is challenged by another factor that may lead to biases, isoprene. Biogenic isoprene is emitted into the troposphere in greater quantities from vegetation than any other volatile organic VOCs [Guenther et al., 1995] and plays a key role in the reactions that form tropospheric O_3 in warmer summer days in certain regions [Chameides et al., 1988; Pierce et al., 1998]. At the same time, CO is also produced from hydroxyl radical (OH) - initiated photo oxidation of isoprene. The co-production of CO and O_3 from isoprene oxidation provides another pathway driving the correlation.

However, the effects of biogenic emissions on the correlation between O_3 and CO and its implications for using the slope approach to estimate anthropogenic O_3 production rate have not been studied in detail. Furthermore, there is few study on the spatial and temporal characters of this contribution. Considering various atmospheric chemical and meteorological conditions in different places, subsequent studies on a greater geographical scope is needed to get insight into the factors affecting tropospheric O_3 fluctuation, which is also favorable for the policy-making in polluted or pollutant exporting regions.

1.1.2 Using the Correlations to Estimate Lower Tropospheric O_3 Concentrations

Monitoring tropospheric O_3 at the large regional and global scales is one of the major tasks of environmental protection. Space-borne remote sensing utilizing its absorption features in the ultraviolet (UV) and the thermal infrared (TIR) bands is the most convenient way to provide O_3 spatial distributions around the globe. However, because of the molecular scattering of UV [Liu et al., 2010] and lack of contrast of TIR [Beer, 2006], the satellite observations for O_3 still show limited sensitivity in the lowermost troposphere, especially near the surface, which is directly relevant to air quality [Cuesta et al, 2013].

3-D air quality model can provide distribution information of surface O_3, but the accuracy of surface O_3 simulations is limited by uncertainties in precursor emissions, atmospheric processes, and nonlinear photochemistry. Some model uncertainties can be mitigated through probabilistic approaches [Dabberdt et al., 2004; Delle Monache et al., 2006a, b; Vautard et al., 2009]. Statistical methods were also applied to estimate O_3 distributions, including classification and regression trees, linear regression, and neural networks [Burrows et al., 1995; Van der Wal and Janssen, 2000; Perez and Reyes, 2006; Cobourn et al., 2007; Shad et al., 2009, Biancofiore et al., 2015]. The advantage of such statistical modeling is that it offers moderate to high accuracy at a moderate cost [Zhang et al., 2012a, b]. However, the nature of statistical modeling often requires a certain number of input variables and does not enable better understanding of chemical and physical processes [e.g., Guillas et al., 2008].

An alternative to numerical or statistical modeling is to make use of the observations of O_3 surrogates. The correlations between O_3 and other chemical species can be used effectively to diagnose O_3 chemical and physical processes [e.g., Parrish et al.,

1993; Chin et al., 1994; Wang and Zeng, 2004; Koo et al., 2012]. CO is often observed and simulated to have a linear relationship with O_3 in the lower atmosphere [e.g., Parrish et al., 1993; Chin et al., 1994; Buhr et al., 1996; Parrish et al., 1998; C´ardenas et al., 1998; Li et al., 2002; Cooper et al., 2002a, b; Honrath et al., 2004; Mao and Talbot, 2004; Huntrieser et al., 2005; Cheng et al., 2017]. Over the eastern United States, the observed slope of O_3 to CO at ~0.3 has contributions from both CO from primary anthropogenic emissions and that from biogenic isoprene oxidation [Cheng et al., 2017] due to the production of both O_3 and CO from biogenic VOCs oxidation [Guenther et al., 1995; Atkinson and Avey, 1998; Pierce et al., 1998; Hudman et al., 2008; Pang et al., 2009; Choi et al., 2010; Geng et al., 2011; Lee et al., 2014; Zhang and Wang, 2016]. It implies that the observations of CO concentrations can be potentially applied to track the contributions to surface O_3 by anthropogenic and biogenic factors. In addition to EPA surface monitoring networks, satellite observations of lower tropospheric CO are more promising than O_3 since CO is more concentrated in boundary layer and unlike O_3 does not have high concentrations in the stratosphere. CO can be detected by satellite TIR, near infrared (NIR) and joint TIR and NIR instruments, such as the NIR SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY (SCIAMACHY) [e.g., De Laat et al., 2012] and the NIR and TIR Pollution in the Measurement of Pollution in the Troposphere (MOPITT) [e.g., Emmons et al., 2004; Straume et al., 2005].

CH_2O is a principal intermediate in the oxidation of atmospheric hydrocarbons in O_3 photo-chemical processes [e.g., Wiedinmyer et al., 2001; Duane et al., 2002; Pang et al., 2009]. It is a major primary radical source leading to ozone production in the presence of nitrogen oxides nitrogen oxides (NO_x) [e.g., Liu et al., 2012]. In addition to direct

emissions from biomass burning and fossil fuel combustion, CH_2O is also produced by oxidation of both biogenic and anthropogenic VOCs. We will show that reginal CH_2O is also correlated to O_3 in section 3.3. Since the sources of CO and CH_2O are often different, the two correlations can provide separate constraints on O_3 distributions. While regulatory monitoring of surface CH_2O is unavailable, CH_2O is detectable from space with good sensitivities in the boundary layer by measuring backscattered solar UV between 325 and 360 nm [Chance et al., 2000], including SCIAMACHY [Wittrock et al., 2006], Ozone Monitoring Instrument (OMI) [Kurosu et al., 2004; González Abad et al., 2015], Global Ozone Monitoring Experiment–2 (GOME-2) [De Smedt et al., 2012], and Ozone Mapping Profiler Suite (OMPS) [Li et al., 2015; González Abad et al., 2016].

1.1.3 Regional Anthropogenic O_3 Precursor Emissions and its Effect on the Correlations

Anthropogenic VOCs are important precursors of ozone and organic aerosols in urban areas. Their sources are poorly quantified in emission inventories as shown by air quality model studies in eastern Texas [Ryerson et al 2003, Parrish et al 2012]. Current emission inventories, based on emission factor calculations, systematically fail to quantify continuous industrial VOCs emissions and that reliable estimates of these emissions can currently only be obtained from measurements. This inventory inaccuracy can finally result in underestimation of CH_2O burden in the air (Figure 1.1). Zhu et al [2014] find the lower CH_2O could be attributed to the fact that instead of direct CH_2O emission, the real emission of VOCs in Houston area should be about 5 times higher than the inventory.

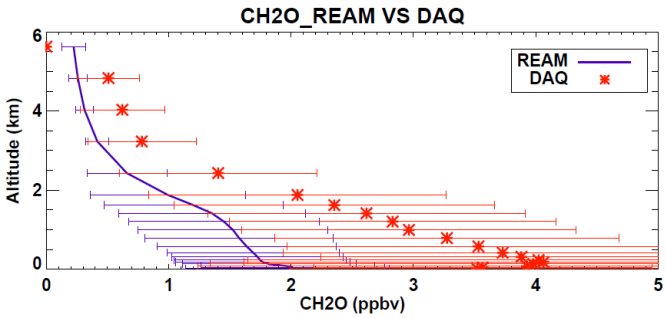

Figure 1.1 – Observed and simulated vertical profiles of CH_2O in 2013.

Several research found that higher alkanes (C>4) plays important role in CH_2O formation [Song et al, 2010; Dasgupta et al, 2005; Johansson et al, 2014]. Johansson et al [2014], with a mobile platform for flux measurements of alkanes using the Solar Occultation Flux (SOF [Mellqvist et al., 2010]) method and mobile differential optical absorption spectroscopy (DOAS [Volkamer et al., 2005]), VOCs, as high as 90% of which is alkanes, are underestimated by a factor of 5-15, due to the missing sources from oil and gas industries in the Houston ship channel (HSC) area.

There are various sources of VOCs emissions in the oil and gas industries with the following major categories: (1) process emissions, including separations, conversions, and treating, (2) combustion emissions for production and transportation purposes, (3) fugitive emissions, including sudden leaks of vapors from equipment or pipelines, as well as continuous small leaks from seals on equipment, (4) auxiliary emissions originate from units like cooling towers, boilers, sulfur recovery units, and wastewater treatment units, and (5) post-processed oil and gas handling, including ship unloading, storage, domestic or ship transportation of oil take place. Extensive emissions tests have been conducted at the areas of oil and gas industries at Houston and HSC [Karl et al., 2003; Ryerson et

al., 2003; Wert et al., 2003; Jobson et al., 2004; De Gouw et al., 2009; Parrish et al., 2009; Mellqvist et al., 2010; Washenfelder et al., 2010]. However, emissions due to ship-loading, storage and transmit of VOCs, a potential problem in every petrochemical industry including any product distribution sites, are currently not addressed neither by the directive on the integrated pollution prevention or by other environmental regulations [Milazzo et al., 2017].

1.2 Description of the Chemical Transport Model

The 3-D Regional chEmical trAnsport Model (REAM) has been applied in a number of tropospheric chemistry and transport studies over North America and East Asia [e.g., Choi et al., 2005, 2008a, 2008b; Jing et al., 2006; Liu et al., 2010, 2012a, 2012b, 2014; Wang et al., 2006, 2007; Yang et al., 2011; Zeng et al., 2003, 2006; Zhao and Wang, 2009; Zhao et al., 2009a, b, 2010; Gu et al., 2013, 2014, 2016; Zhang et al., 2016b; Zhang and Wang, 2016; Zhang et al., 2017]. The model has two different horizontal resolution of 36×36 km^2 and 4×4 km^2. Transport is driven by the Weather Research and Forecasting (WRF) model assimilated meteorological fields constrained by the Climate Forecast System Reanalysis (CFSR) products (http://cfs.ncep.noaa.gov/cfsr). The WRF model domain is 10 grid-cell larger than that of REAM on each lateral side. The chemistry mechanism in REAM is adopted from the GEOS-Chem model (V9-02) [Bey et al., 2001] with updates of kinetics data (http://jpldataeval.jpl.nasa.gov). The anthropogenic emission inventory used in the model is the 2011 National Emission Inventory (NEI) (https://www.epa.gov/air-emissions-inventories/2011-national-emissions-inventory-nei-

data). The biogenic isoprene emissions are computed using the Model of Emissions of Gases and Aerosols from Nature (MEGAN) version 2.1 [Guenther et al., 2012]. For 36 × 36 km^2 resolution model, initial and boundary conditions for chemical tracers are taken from the GEOS-Chem (V9-02) 2° × 2.5° simulation results [Bey et al., 2001]. For 4 × 4 km^2 resolution model, initial and boundary conditions are taken from the results of 36 × 36 km^2 resolution model with the same time period.

CHAPTER 2. ANTHROPOGENIC AND BIOGENIC CONTRIBUTIONS TO BOUNDARY LAYER O_3-CO REGRESSION SLOPE IN SUMMER

2.1 Introduction

Tropospheric O_3 and CO are major pollutants in the troposphere [e.g., Logan et al., 1981; Wang and Jacob, 1998; Wang et al., 1998; Lelieveld and Dentener, 2000]. The ambient O_3 and carbon monoxide CO burdens depend on a wide range of physical and chemical processes that take place in the troposphere [Logan et al., 1981; Wang and Jacob, 1998; Wang et al., 1998; Lelieveld and Dentener, 2000]. Both of these two species exert influence on the tropospheric oxidizing capacity and affecting the air quality and climate. The tropospheric lifetime of O_3 ranges from days to months [Wang et al., 1998], and that of CO varies from weeks to months [Duncan et al., 2007]. Their relatively long lifetimes enable these two species to undergo intercontinental transport which exerts upon the chemistry of the troposphere, especially with regard to the ozone budget.

A linear relationship between the two pollutants is often observed and simulated in the lower atmosphere [e.g., Parrish et al., 1993; Chin et al., 1994; Buhr et al., 1996; Parrish et al., 1998; C´ardenas et al., 1998; Li et al., 2002; Cooper et al., 2002a, b; Honrath et al., 2004; Mao and Talbot, 2004; Huntrieser et al., 2005]. Aircraft measurements, employing northeastern United States and the pollutant outflow region in the northern Atlantic as the most prevailed studied areas, showed that O_3-CO correlations in the summer are positive and often strong, depending on the weather patterns during the measurement period [Buhr et al., 1996; Cooper et al., 2002a, b; Huntrieser et al., 2005]. One of the reasons of this correlation is the synchronous emissions of anthropogenic CO and NO_x which is one of the

most important precursors of O_3. Since CO has long been used as one of typical anthropogenic tracers [Khalil and Rasmussen, 1994; Holloway et al., 2000; Seinfeld and Pandis, 2016; Duncan et al., 2007], the observed slope of O_3 to CO at ~0.3 has been used to estimate net anthropogenic O_3 production [e.g., Parrish et al., 1993; Chin et al., 1994]. However, these and subsequent studies also noted the uncertainty in using the O_3-CO slope approach. For example, O_3 deposition will lead to slopes < 0.3 in the region of anthropogenic emission area [Chin et al., 1994; Parrish et al., 1998; C´ardenas et al., 1998; Li et al., 2002; Mao and Talbot, 2004; Real et al., 2008] and stratospheric intrusion will affect the slopes at high-altitude regions [Parrish et al., 1998].

Another factor that may lead to biases using this approach is the contribution of biogenic emissions to CO production. Chin et al. [1994] suggested that biogenic isoprene oxidation might be a very important process. Biogenic isoprene is emitted into the troposphere in greater quantities from vegetation than any other VOCs [Guenther et al., 1995] and plays a key role in the reactions that form tropospheric O_3 in warmer summer days in certain regions [Chameides et al., 1988; Pierce et al., 1998]. In regions of large biogenic emissions, oxidation of isoprene far exceeds oxidation of anthropogenic VOCs, leading to production of both O_3 and CO [Guenther et al., 1995; Atkinson and Avey, 1998; Pierce et al., 1998; Hudman et al., 2008; Pang et al., 2009; Choi et al., 2010; Geng et al., 2011; Lee et al., 2014; Zhang and Wang, 2016]. As anthropogenic emissions decrease in the U.S. [e.g., Parrish et al., 2002; Hassler et al., 2016], the relative importance of biogenic emissions increases.

Extensive sampling of O_3, CO, and other trace gases was conducted during the NASA Earth Venture airborne campaign, Deriving Information on Surface Conditions from Column and Vertically Resolved Observations Relevant to Air Quality (DISCOVER-

AQ), in July 2011 around the Baltimore-Washington area (see the online data archive: http://www-air.larc.nasa.gov/missions/discover-aq/discover-aq.html for detailed description of the research flights). In this study, we apply a 3-D chemical transport model to quantitatively study factors contributing to the observed O_3-CO correlation and regression slope and examine the implications for using the observed correlation.

2.2 Data and Methods

REAM model has a horizontal resolution of 36 × 36 km^2. Transport is driven by the WRF model assimilated meteorological fields constrained by the CFSR products (http://cfs.ncep.noaa.gov/cfsr). The WRF model domain is 10 grid-cell larger than that of REAM on each lateral side. The chemistry mechanism in REAM is adopted from the GEOS-Chem model (V9-02) [Bey et al., 2001] with updates of kinetics data (http://jpldataeval.jpl.nasa.gov). The anthropogenic emission inventory used in the model is the 2011 NEI (https://www.epa.gov/air-emissions-inventories/2011-national-emissions-inventory-nei-data). The biogenic isoprene emissions are computed using the MEGAN version 2.1 [Guenther et al., 2012]. Initial and boundary conditions for chemical tracers are taken from the GEOS-Chem (V9-02) 2° × 2.5° simulation results [Bey et al., 2001].

The observational data were obtained during the NASA 2011 DISCOVER-AQ airborne campaign, sampling from Washington's Beltway northeast to Baltimore and continuing on to the Delaware State line and occasionally over the Chesapeake Bay. The vertical structures of pollutants were measured through 253 daytime vertical profiles from 300 m to 5 km over six selected locations during 14 flights by NASA's P-3B aircraft between June 27 and July 31. CO was measured by a diode laser spectrometer [Sachse et al., 1987] with a 2% uncertainty. NO, NO_2, and O_3 were measured by the National Center for Atmospheric Research 4-channel chemiluminescence instrument [Brent et al., 2015]

with 10%, 15% and 5% uncertainties, respectively. CH_2O was measured by a difference frequency generation absorption spectrometer [Weibring et al., 2010] with a 4% uncertainty. Isoprene was measured by a proton-transfer-reaction mass spectrometer [Lindinger et al., 1998] with a 10% uncertainty.

The model domain of REAM and locations of DISCOVER-AQ observations are shown in Figure 2.1. There are 253 aircraft spirals around 6 locations [Zhang et al., 2016b]. To evaluate model simulations with the observations, we first identify the model grid cells corresponding to the locations of aircraft spirals. These model profiles are then archived at the time of aircraft sampling. The corresponding aircraft and model data are used in correlation analysis and the averaged model vertical profiles are also compared to the observations.

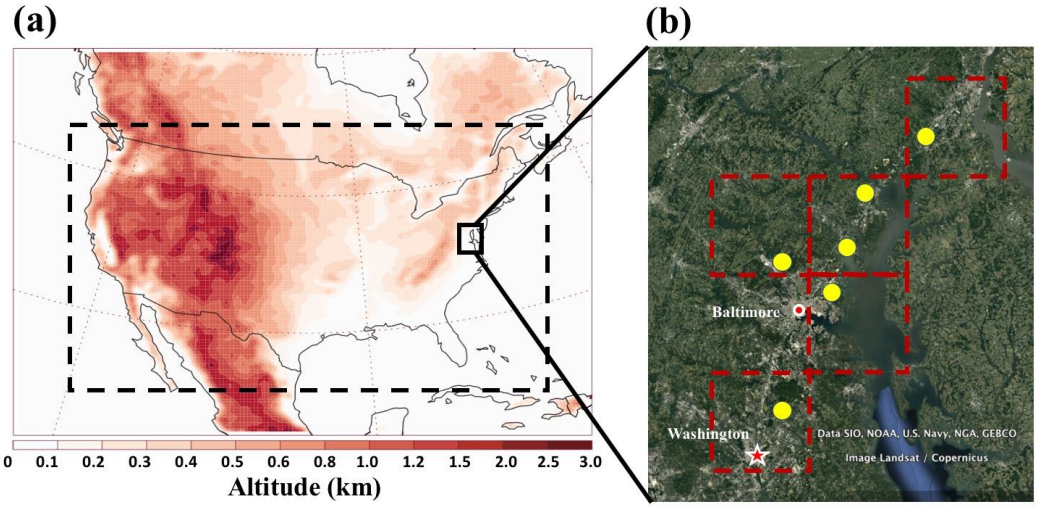

Figure 2.1 – Overview of regions involved in this study. The black dashed rectangle in (a) represents the model domain of REAM. The solid black rectangle in (a) represents the domain of DISCOVER-AQ observations. The locations of 6 aircraft observation spiral sites are shown by the yellow round dots in (b). The red dashed rectangles in (b) represent the model cells, where these spiral locations reside in REAM.

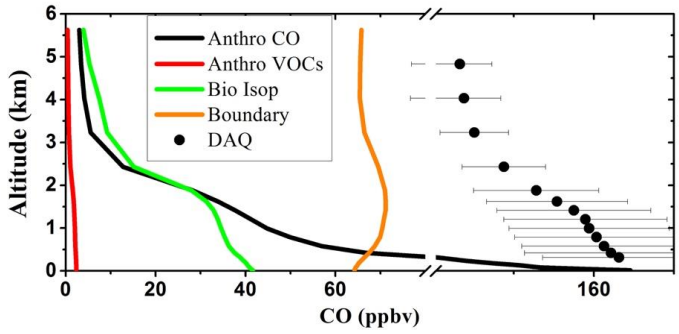

Figure 2.2 – Simulated daytime (10 A.M. to 6 P.M.) vertical profiles of CO tracers. "Anthro CO", "Anthro VOCs", "Bio ISOP", and "Boundary", denote CO due to primary anthropogenic emissions, anthropogenic VOC oxidation, biogenic isoprene oxidation, and transport from the model boundary, respectively. "DAQ" denotes DISCOVER-AQ observations.

In order to analyze the attributions of the observed O_3-CO relationship, we trace separately four different CO sources: primary anthropogenic emissions ($CO_{anthroCO}$), the oxidation of anthropogenic VOCs ($CO_{anthroVOCs}$), the oxidation of biogenic isoprene ($CO_{bioISOP}$), and transport from model lateral and upper boundaries (CO_{BC}). We do not account for other biogenic VOCs since isoprene is much more abundant than the other species providing the source for the vast majority biogenic CO [e.g., Kesselmeier and Staudt, 1999; Lathière et al., 2006; Guenther et al., 2012; Sindelarova et al., 2014]. In tagged-tracer simulations, relevant species and radicals, such as O_3, NO_x, and HO_x (OH and HO_2), are fixed using results archived from the standard simulation. The sum of the four individual tracers is within 5% of the simulated total CO concentrations in the standard simulation for grid cells over the Washington-Baltimore region. Minor scaling adjustments, assuming that relative CO attributions stay the same, are added in post-processing to ensure that the sum of the CO tracers is the same as the total CO for each grid cell in the standard

simulation. With simulated CO attribution results, we can decompose the O_3-CO regression slope into 4 sub-slopes of the corresponding CO tracers (Eq. (1), see the Appendix A1)

$$\text{Least} - \text{Squares Regression Slope} = \frac{Cov(\mathrm{CO_{anthroCO}}, \mathrm{O_3})}{Var(\mathrm{CO_{total}})} + \frac{Cov(\mathrm{CO_{anthroVOCs}}, \mathrm{O_3})}{Var(\mathrm{CO_{total}})} + \frac{Cov(\mathrm{CO_{BC}}, \mathrm{O_3})}{Var(\mathrm{CO_{total}})} + \frac{Cov(\mathrm{CO_{bioISOP}}, \mathrm{O_3})}{Var(\mathrm{CO_{total}})} \quad (1)$$

where *Cov* and *Var* denote covariance and variance, respectively. Eq. (1) shows that the contributions of each CO tracer to the O_3-CO regression slope is proportional to its covarince with O_3. It is therefore possible to have both positive and negative slope contributions.

2.3 Biogenic Contribution to Slope of O_3-CO at Maryland

2.3.1 Comparison of Tracers Concentrations between Simulations and Observations

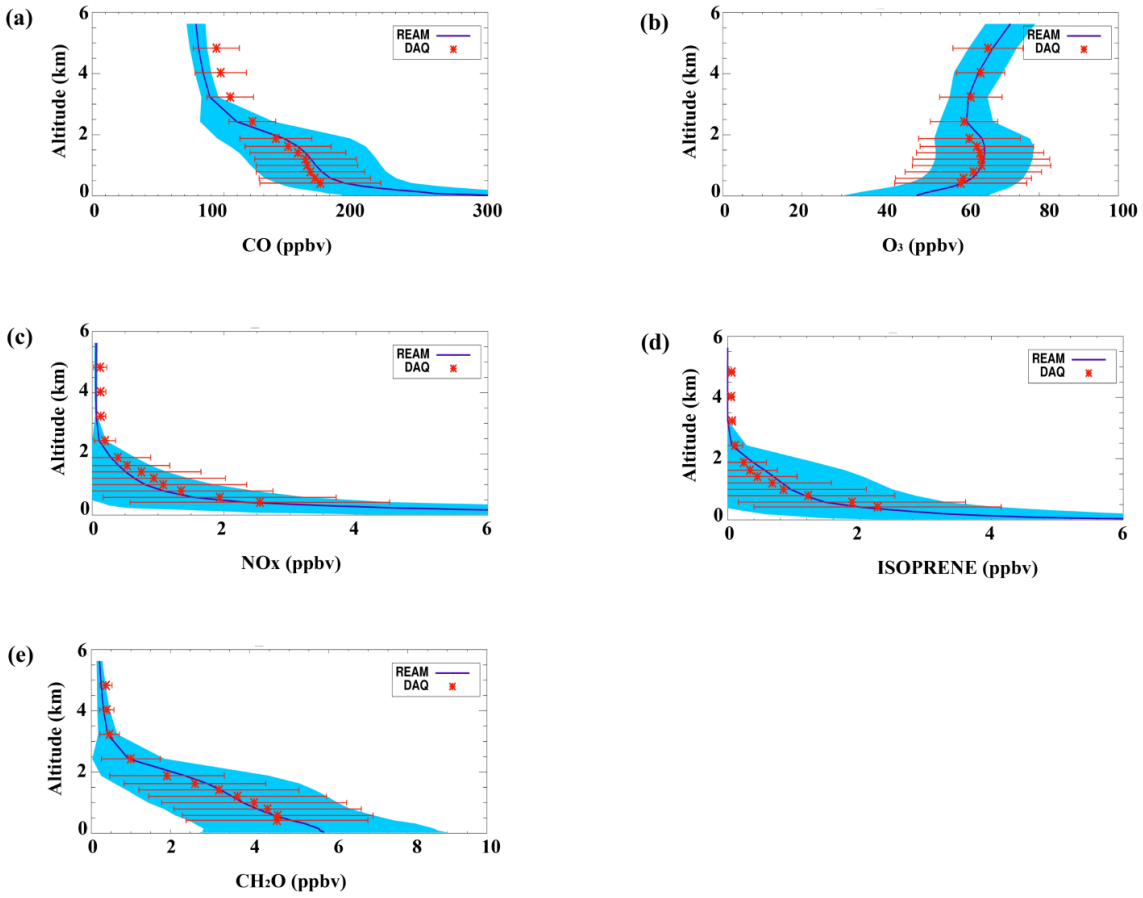

Figure 2.3 – Observed and simulated vertical profiles of CO (a), O_3 (b), NO_x (c), isoprene (d) and CH_2O (e) in July 2011 during the DISCOVER-AQ experiment. DAQ represents DISCOVER-AQ data. The red horizontal bars show the observed standard deviations; the shaded blue areas denote the simulated standard deviations. For comparison purposes, the observational data are binned vertically according to the vertical grid structure in the model.

We first evaluate simulated vertical profiles of O_3, CO, NO_x, isoprene, and CH_2O with DISCOVER-AQ observations (Figure 2.3). CH_2O is an intermediate product of anthropogenic and biogenic VOCs oxidation; NO_x is mainly from anthropogenic emissions and isoprene is biogenic, both of which contribute to O_3 production [e.g., Zhang et al., 2016a]. The model reproduces well the observed vertical profiles of these species.

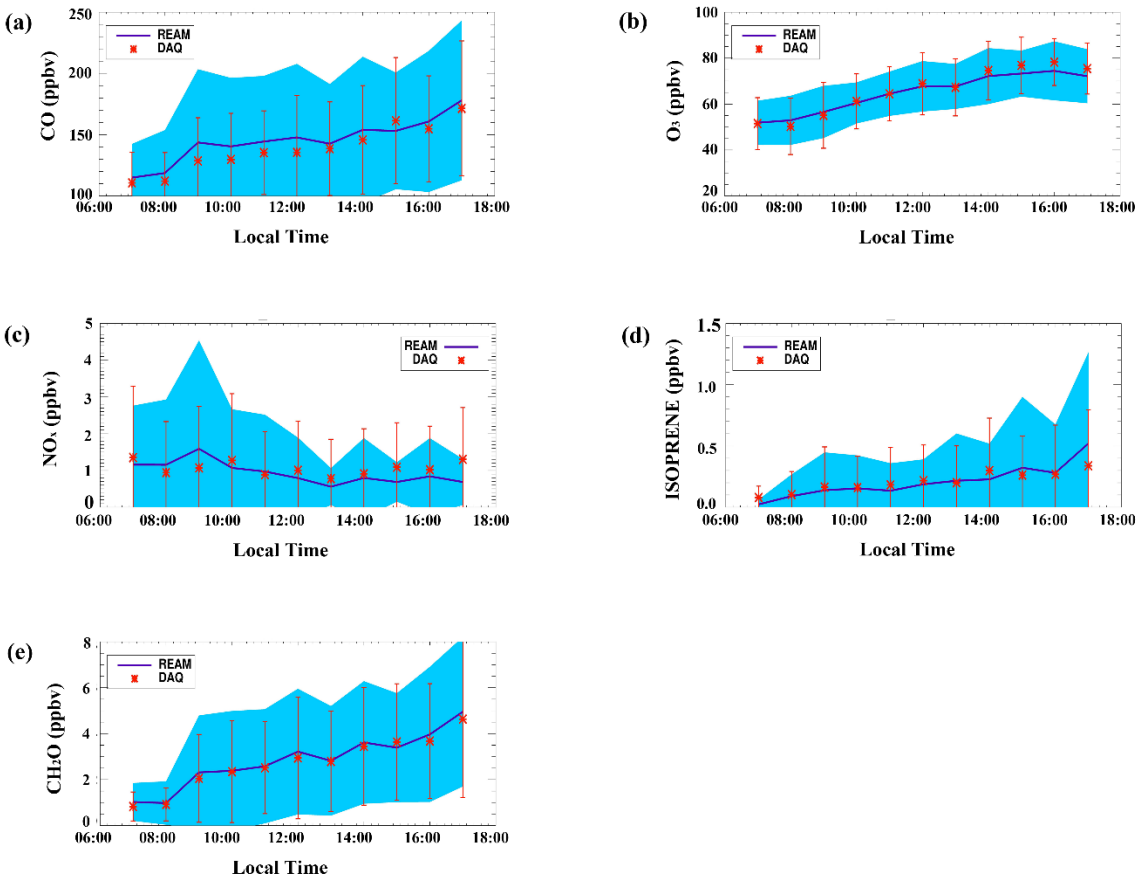

Figure 2.4 – Same as Figure 2.3 but for hourly trace gases mixing ratios. The red vertical bars show the observed standard deviations; the shaded blue areaes denote the simulated standard deviations. Observed or corresponding simulated data from 300 m to 2.5 km are averaged hourly.

The short-lived NO_x, isoprene, and CH_2O decrease rapidly from the surface to the free troposphere, reflecting in part the dominant NO_x and VOCs sources near the surface and rapid photochemical loss in the atmosphere. The nearly linear decrease of CH_2O, in contrast to the exponential decreases of NO_x and isoprene, reflects that its atmospheric secondary (photochemical) source is much larger than primary emissions. Ozone has a peak in the middle of the boundary layer due in part to dry deposition loss of ozone and in

part to decreasing photochemical production and loss with altitude. It increases with altitude in the free troposphere reflecting the increase of chemical lifetime with altitude.

Using results from tagged-tracer simulation, we can attribute the vertical profile of CO to 4 different sources (Figure 2.2). The boundary contribution does not vary significantly with altitude. The contribution by primary emissions from the surface decreases with altitude. The contributions by secondary productions from biogenic or anthropogenic VOCs do not have as large a decrease as those of primary emissions since secondary production takes place mostly from late morning to the afternoon when the boundary layer is well mixed, while the shallow boundary layer in early morning (as early as 5 am) can lead to a large gradient of surface emitted species such as CO [e.g., Zhang et al., 2016]. The rapid decrease of secondary CO from the boundary layer to the free troposphere reflects slow vertical transport between the two layers. Clearly secondary CO is a major factor shaping its observed vertical variation.

The daytime hourly variations of the selected species averaged from 300 m to 2.5 km are shown in Figure 2.4. In general, the variation is much less than in the vertical. The temporal change of NO_x is insignificant while O_3, CO, isoprene, and CH_2O tend to accumulate in the lower atmosphere from morning to afternoon. The observed variations are well simulated by the model, providing further evidence for the good model performance during the measurement period.

2.3.2 O_3-CO Correlations

Correlations between chemical species can be used effectively to diagnose atmospheric processes [e.g., Parrish et al., 1993; Chin et al., 1994; Wang and Zeng, 2004; Koo et al., 2012]. We compare simulated O_3-CO correlations to the observations as a

function of altitude or time of the day in Figure 2.5. The observations show a narrow range of R value (~0.75) from 1 km to the top of the boundary layer (~2.5 km). Below 1 km, the observations indicate a trend of decreasing R value with altitude down to 300 m. The model results are similar to the observed values and also show lower correlations near the surface and in the free troposphere. In the free troposphere, CO is dominated by the contribution by boundary layer transport (Figure 2.2), which is not related to O3 production in the boundary layer. Near the surface, CO is dominated by surface primary emissions, which are not directly related to O_3 production either. Furthermore, mixing of fresh emissions with photochemically aged air in the boundary layer leads to an increase of O_3 but a decrease of CO since the large vertical gradients of O_3 and CO near the surface have opposite signs (Figure 2.2). In the middle and upper boundary layer, the relative contribution by photochemical CO production increases. Therefore, the concurrent photochemical production of O_3 and CO is likely the major factor contributing to the observed positive correlation between O_3 and CO. We quantify this contribution by decomposing the O_3-CO regression slope into the different CO sources.

Figure 2.5 also shows the observed O_3-CO correlation coefficient as a function of time of the day with data from 300 m to 2.5 km. This R value represents the spatial correlation in a given hour. It is somewhat lower than the temporal correlation discussed previously. The value ranges from 0.5 to 0.6 in the morning and 0.6 to 0.8 in most of the afternoon. The model is in good agreement with the observations except the overestimation at 5 pm.

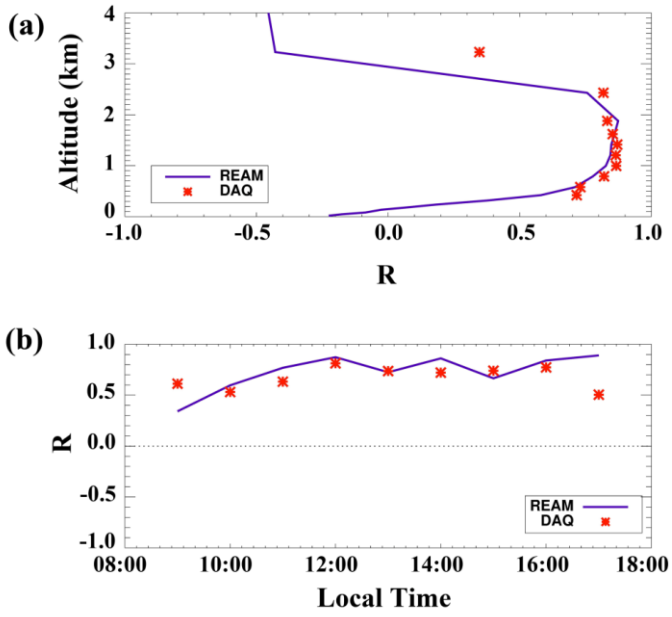

Figure 2.5 – Observed and simulated O_3-CO correlation coefficients (R) as a function of altitude for daytime of 10 am - 4 pm (a) and local time for altitude of 0.3 - 2.5 km (b). The R values are computed using the observation or corresponding model data (for the entire DISCOVER-AQ experiment) at a given altitude bin or a given time period.

2.3.3 Regression Slopes and Sub-slopes of O_3-CO

Figure 2.6 compares the observed and simulated O_3-CO regression slope. The O_3-CO regression slope at a given altitude is underestimated by the model at 1-2.5 km by ~20%, but is overestimated below 1 km. At 3 km, the model shows a negative regression slope value opposite to the observations. However, the observed decrease of the regression slope below 1 km and above 2.5 km is captured by the model. The observed O_3-CO regression slope at a given hour is better simulated than that at a given altitude. The O_3-CO regression slope of 0.1-0.3 during the DISCOVER-AQ experiment is consistent with the previous reports for land areas [e.g., Chin et al., 1994]. We apply the model results to understand the relative contributions of CO sources to the O_3-CO regression slope (Eq. (1)).

In the lower atmosphere (below 2.5 km), where the model results are in reasonably good agreement with the observations, Figure 2.6 shows that the contribution of CO from biogenic isoprene oxidation is nearly as large as that of primary anthropogenic CO. In particular, the relative importance of biogenic CO increases towards the surface. These two contributions account for > 90% of the O_3-CO regression slope in the lower atmosphere (< 2.5 km). The contributions of CO from photochemical oxidation of anthropogenic VOCs and from the model boundaries are minor in comparison. In the free troposphere (above 2.5 km), Figure 2.6 shows that the contributions by anthropogenic CO and VOCs and biogenic isoprene to the O_3-CO regression slope are negative. This is because air near the top boundary layer is enriched in these CO tracers but not in O_3 relative to low free tropospheric air (Figure 2.3). The positive contribution by CO transported from the (lateral) boundary reflects largely a general south-to-north positive gradient in free tropospheric CO and O_3 in the model.

Our analysis of the DISCOVER-AQ measurements suggests that the previously used hypothesis, i.e., the regression slope of O_3 to CO was controlled by anthropogenic emissions and the photochemical oxidation of anthropogenic ozone precursors [e.g., Parrish et al., 1993; Chin et al., 1994], is no longer valid. However, a caveat is that anthropogenic emissions of CO decreased by ~46% from 1993 to 2011 (https://www.epa.gov/air-emissions-inventories/air-pollutant-emissions-trends-data). The large decrease of anthropogenic CO emissions can lead to the invalidity of the hypothesis. We carry out a sensitivity simulation in which the anthropogenic primary CO emissions are doubled (Figure 2.7). After doubling anthropogenic primary CO emissions, anthropogenic CO accounts for a larger fraction in the lower troposphere. More drastic

change is found in the O_3-CO regression slope, ~75% of which in the lower troposphere is due to anthropogenic emissions, making the previous hypothesis a good approximation. As anthropogenic emissions continue decreasing, we expect that the biogenic contribution to the O_3-CO regression slope will increase in the future.

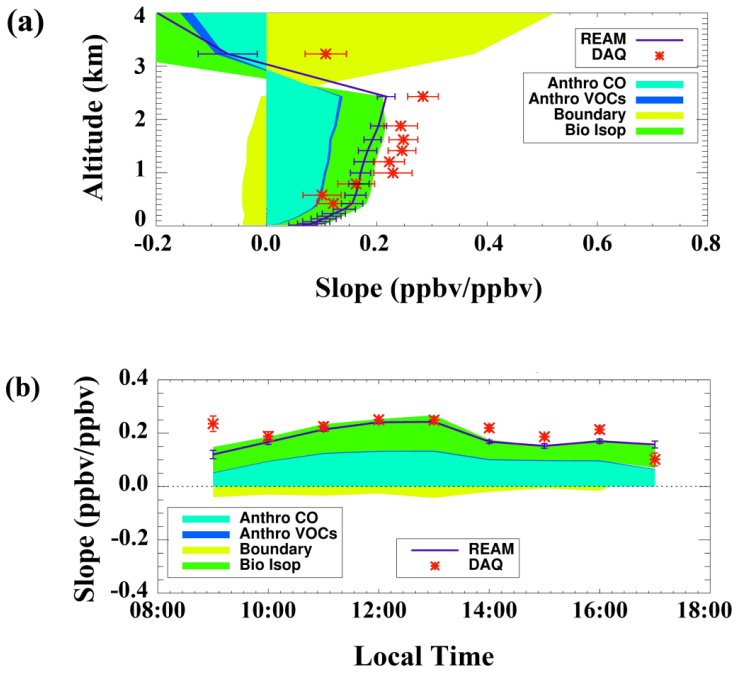

Figure 2.6 – Same as Figure 2.5 but for the observed and simulated vertical profiles and diurnal variations of the CO-O_3 regression slope and sub-slopes. The components of sub-slopes, due to varied CO sources, are shown using areas filled with different colors. The legends for different CO sources are the same as in Eq. (1). The horizontal bars in (a) and vertical bars in (b) show the observed or simulated standard deviation of the regression slope. The slope and sub-slopes values are computed using the observations or model data at a given altitude bin or for a given time period.

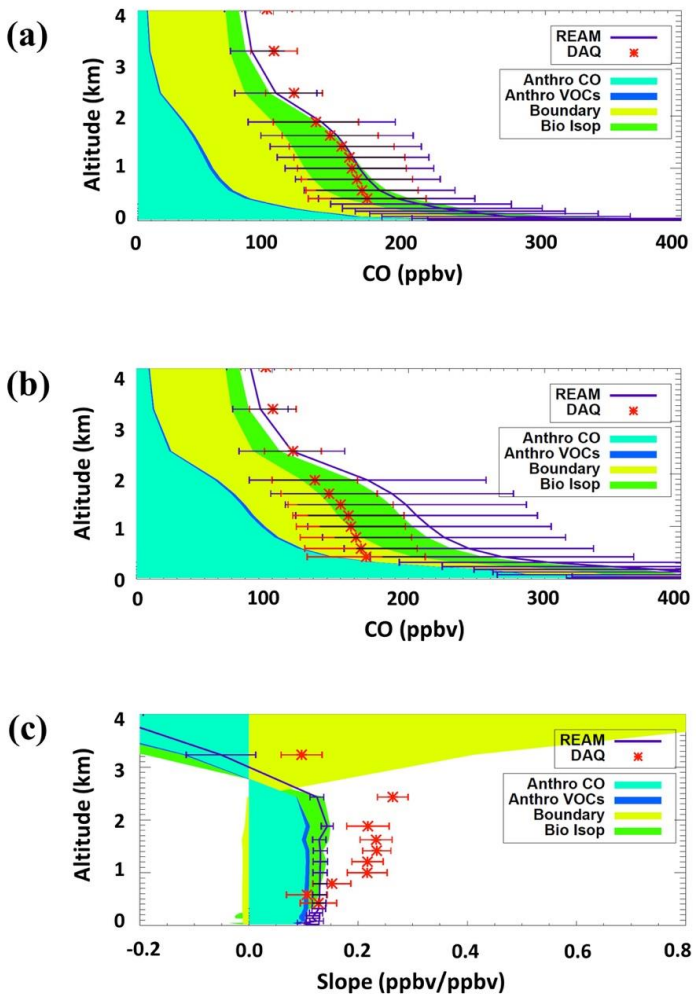

Figure 2.7 – Sensitivity analysis of regression slope decomposition. (a) Simulated vertical CO profiles from four different sources with original emissions. (b) Same as (a) but with doubled primary anthropogenic CO emissions. (c) Same as Figure 3 (a) but with doubled primary anthropogenic CO emissions.

2.4 Biogenic and Anthropogenic Contributions to Slopes of O_3-CO at United States and China

In this study, we apply a 3-D chemical transport model to quantitatively study the O_3-CO regression slope and its sub-slopes (Eq. (1)) in the whole regions of United States

and China to evaluate the contributions from both anthropogenic emissions and biogenic isoprene photo-chemistry.

REAM model setup in this study is the same with that in chapter 2. Model simulation starts from July 1st to July 31nd, 2011. Data is selected hourly from 10:00 A.M. to 4:00 P.M. in local daytime and grouped to calculate the correlations for each grid box.

2.4.1 Distribution of Correlations of O_3-CO in United States

Shown in the standard slope distribution in Figure 2.8, total O_3-CO slope that ranges from 0.2 to 0.5 presents in 3 major outflow regions in oceans: North Atlantic oceans area, Gulf of Mexico (towards the Southeast) and Eastern Pacific (towards the Southwest). This is fairly consistent with the observation located at North Atlantic oceans area [Chin et al., 1994; Zhang et al., 2006; Voulgarakis et al., 2011]. The homogeneous slope value in the ocean is a result of long life time of CO and aged O_3 in absence of surface emissions. Since the southeastern United States is not only one of the largest biogenic isoprene emission areas but also an important pollutant provider onto the pollutant outflow region in the northern Atlantic, an expected strong contribution from biogenic isoprene oxidation presents in the regions of North Atlantic and the Mexico Gulf. The contribution from anthropogenic primary CO emission and biogenic isoprene oxidation to the total O_3-CO slope in these two regions are both within the range of 0.1 to 0.2, indicating their approximately equal contributions to the total slope. This is apparently different from the previous researches' assumption that O_3-CO slope was majorly anthropogenic process controlled in the pollutant outflow region of North Atlantic [Chin et al., 1994; Voulgarakis et al., 2011; Zhang et al., 2006]. However, the scenario is different at the outflow region of western United States. The air mass on the ocean northeast downwind from the west coast shows that primary anthropogenic CO emissions contribute the most of total slope, but biogenic isoprene oxidation does a little.

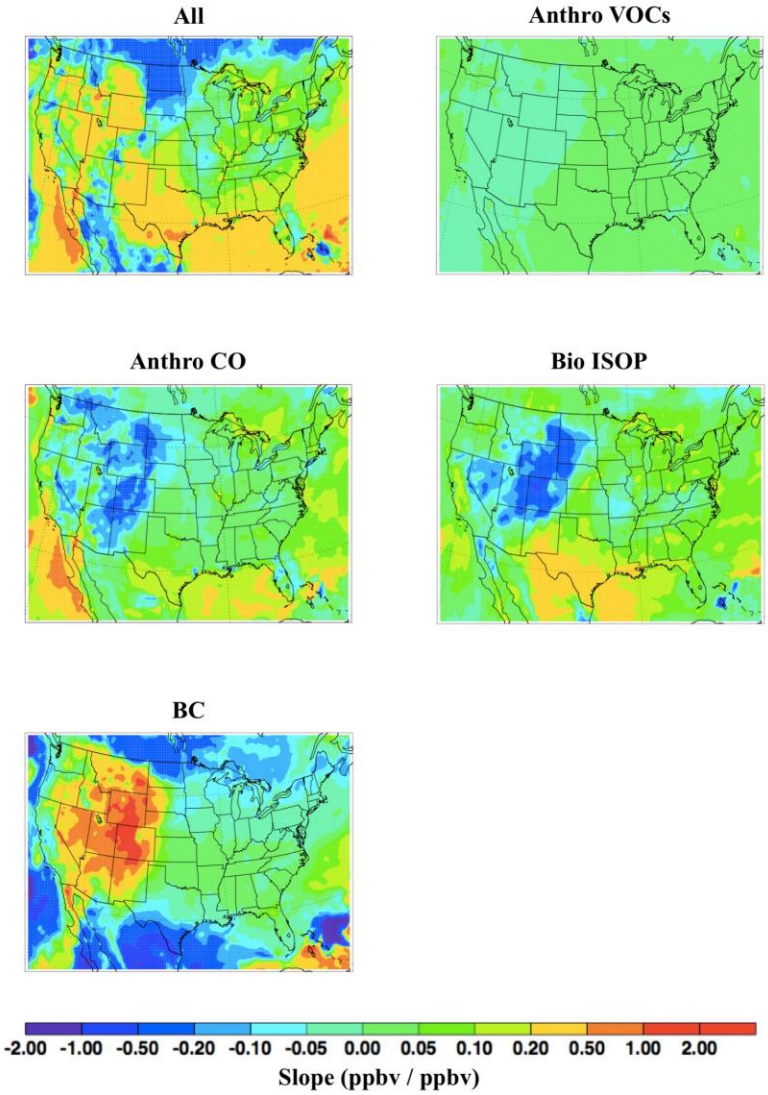

Figure 2.8 – Distribution of slope ($\Delta O_3/\Delta CO$) and sub slopes in surface level of United States. The title "Standard" stands for the slope computed by total O_3 and total CO concentration; the titles of "Anthro VOCs", "Bio ISOP", "Anthro VOCs", and "BC" stand for the sub slopes computed by total O_3 and CO from primary anthropogenic emissions, oxidation of anthropogenic VOCs, oxidation of biogenic isoprene, and transport from the lateral and upper model boundaries, respectively.

On the land surface, the components of sub-slopes are different between western part and eastern or mid-eastern parts of United States, disputing the background areas along

the Rocky Mountain. Lower values of O_3-CO slope (0-0.2) widely distribute at the land surface in eastern United States. This is mainly attributed to the lower contribution by primary anthropogenic CO emissions. Nevertheless, biogenic isoprene oxidation acts as the dominant contributor to the total slope which is consistent with the scenario that the relative importance of biogenic CO increases towards the surface in Washington-Baltimore area [Cheng et al., 2017]. On the contrary, the contribution by primary anthropogenic CO emissions in western United States is slightly more than biogenic isoprene oxidation.

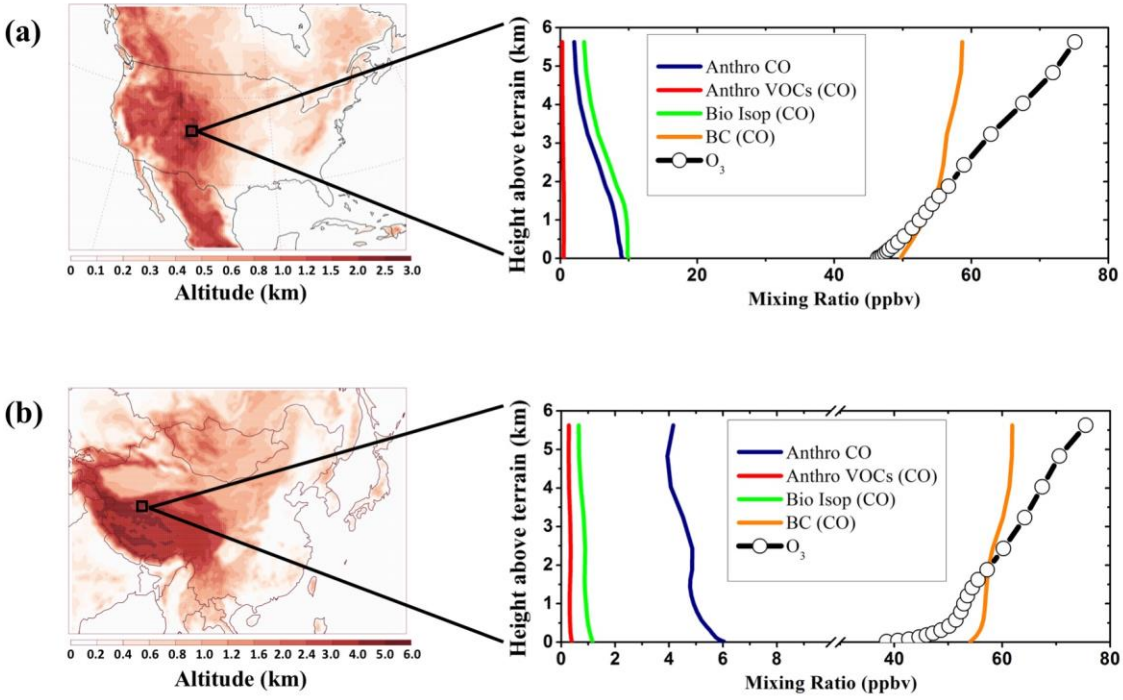

Figure 2.9 – Simulated daytime (10 A.M. to 3 P.M.) vertical profiles of O_3 and 4 different traced CO in high altitude background regions. "Anthro CO", "Anthro VOCs", "Bio ISOP" , and "BC", denote CO due to primary anthropogenic emissions, anthropogenic VOCs oxidation, biogenic isoprene oxidation, and transport from the model boundary, respectively. "O_3" denotes the total O_3 concentrations.

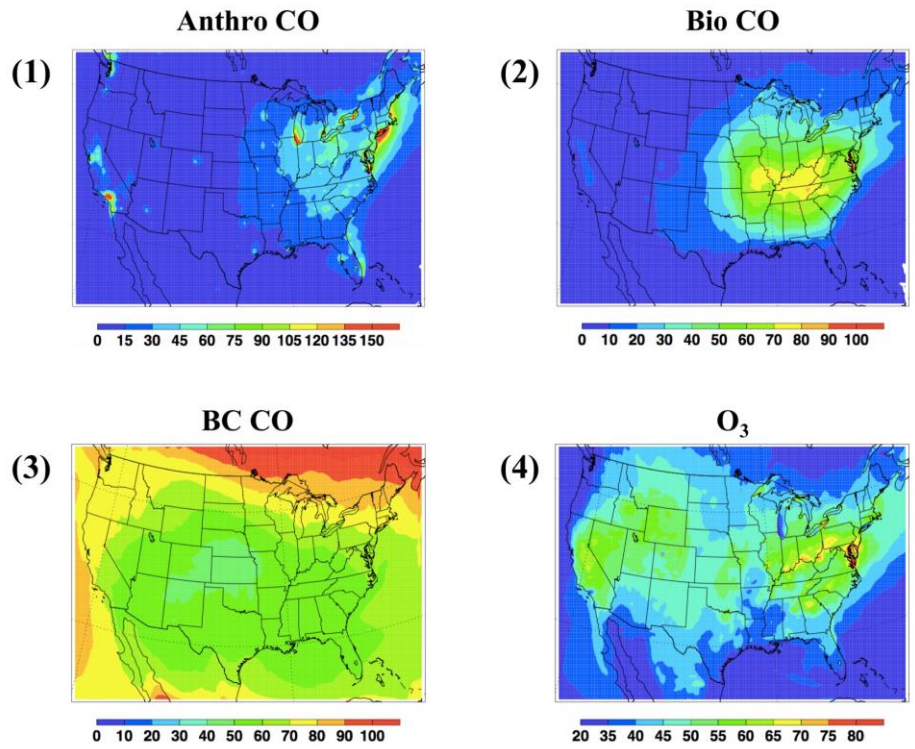

Figure 2.10 – Simulated daytime (10 A.M. to 3 P.M.) distributions of O_3 and 3 different traced CO near surface in United States: CO due to (a) primary anthropogenic emissions, (b) biogenic isoprene oxidation, and (c) transport from the model boundary; and (d) total O_3 concentration.

The importance of contribution by CO transported from the (lateral) boundary to the O_3-CO regression slope is more significant at background areas. This contribution is positive at the rocky mountain in the mid-western of United States. It is because the mixing of the surface air with the air above it leads to an increase of both O_3 and CO from boundary since the vertical gradients of O_3 and CO from boundary have same signs near surface in the higher altitude areas (Figure 2.9). Correspondingly, the contributions by anthropogenic CO and biogenic isoprene are negative since the vertical gradients of these CO tracers have opposite signs to that of O_3 (Figure 2.9). On the contrary, the scenario is opposite near the

lateral boundaries in south Canada, North Mexico and surrounding oceans. It is because in the lower free troposphere in these areas, the boundary-to-interior gradient of O_3 is same signed to that of CO transported from the lateral boundary but opposite signed to other tagged CO tracers (Figure 2.10).

Correlations coefficient (R value) between chemical species can be used effectively to diagnose atmospheric processes [e.g., Parrish et al., 1993; Chin et al., 1994; Wang and Zeng, 2004; Koo et al., 2012]. We also compute the distribution of R values of O_3 and four different traced CO (Figure 2.11) to verify the reasonability of the slope and sub-slope discussed above. Strong correlations present on the outflow oceans (R is ~0.75 to 0.9 in standard simulation; R is ~0.5 to 0.75 for O_3 with CO from primary anthropogenic emissions, oxidation of biogenic isoprene, and oxidation of anthropogenic VOCs). And modest correlations present on the continent (R is ~0.5). The patterns of distributions of R values are very similar to those of corresponding slope and sub-slopes of O_3 and traced CO, respectively, except for sub-slopes of O_3 and CO from oxidation of anthropogenic VOCs. It should be noted that even the correlations of O_3 and CO from anthropogenic VOCs oxidations is strong, the contribution from this traced CO is minor throughout the whole domain due to its small amount.

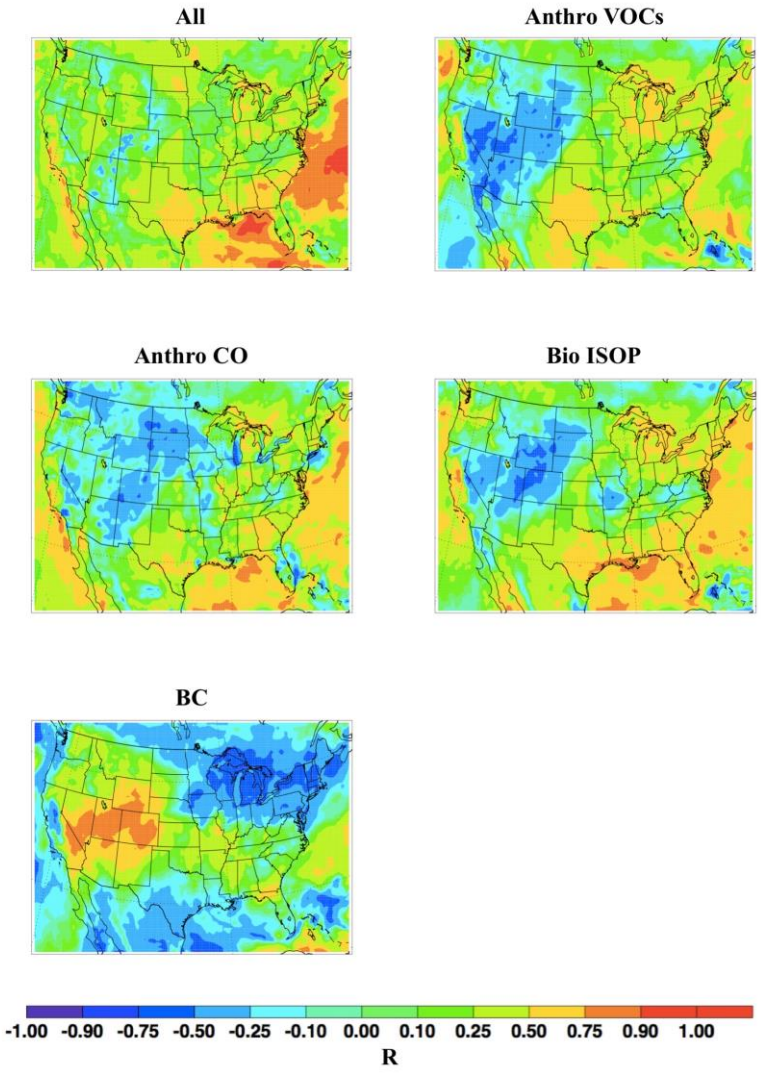

Figure 2.11 – Distribution of correlation coefficient (R) of O_3-CO near surface of United States. The title "Standard" stands for the R computed by total O_3 and total CO concentration; the titles of "Anthro VOCs", "Bio ISOP", "Anthro VOCs", and "BC" stand for the R computed by total O_3 and CO from primary anthropogenic emissions, oxidation of anthropogenic VOCs, oxidation of biogenic isoprene, and transport from the lateral and upper model boundaries, respectively.

Apparent difference of regional O_3-CO slope decompositions between eastern and western parts of United States is shown in Figure 2.12. For the eastern part, even in the urban areas with high anthropogenic emissions like New York (c), Baltimore (f) and

Atlanta (h), the contribution from biogenic isoprene oxidation takes place more than half proportion over all the compositions. Correspondingly, in the pollutant export regions of eastern oceans (e, g, i) and Seal Island (d) passed by the continent outflows, this contribution makes as approximately equivalent contribution as primary anthropogenic CO does. On the contrary, at western coasts of United States like in Los Angeles (k), CO from biogenic isoprene oxidations is negligible contrasted to the significant primary anthropogenic emissions of CO (Figure 2.10), which results in an overwhelming contribution from the later than the former regionally and in the outflow region in Pacific (j). Compared with other mountain areas (a) that are boundary transport controlled, Denver (b), with high anthropogenic emissions, also shows a considerable negative contribution by primary anthropogenic CO emissions. Thus, the total slope in Denver is negative and in small magnitude due to the sub-slopes mutual offset.

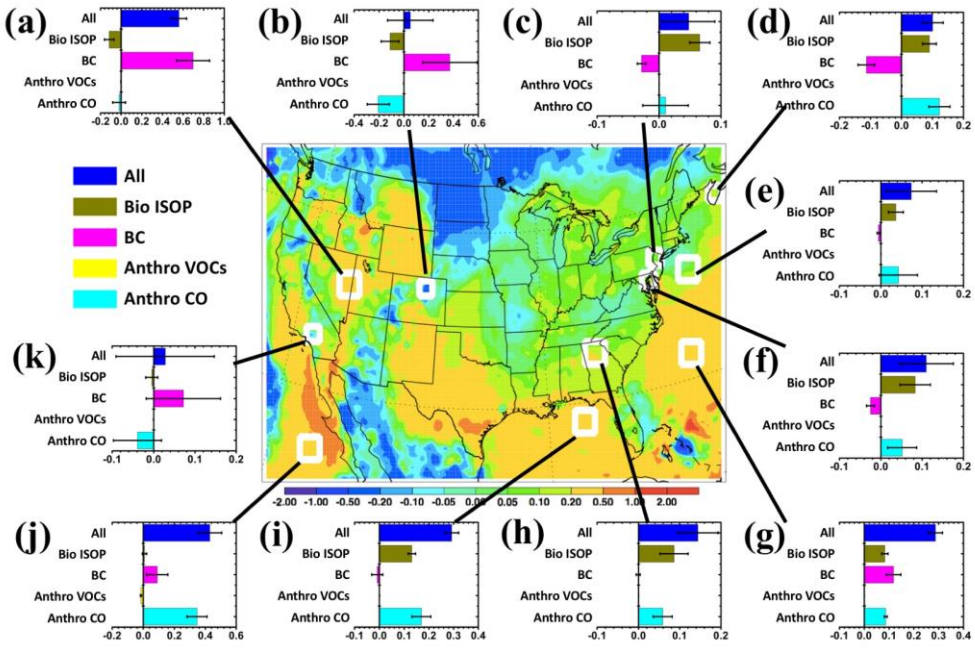

Figure 2.12 – Regional compositions of O_3-CO slope and sub slopes in United States: (a) Mountain Region, (b) Denver, (c) New York, (d) Seal Island, (e, g, i) North Pacific

ocean regions, (f) Baltimore, (h) Atlanta, (j) Eastern Pacific ocean areas, (k) Los Angeles.

2.4.2 Distribution of Correlations of O_3-CO in China

Distributions of O_3-CO slope and sub slopes in China region at the same period of interest (July 2011) are shown in Figure 2.13. The total slope distribution (All) shows a similar scenario in the China ocean areas to that of United States. In most areas of East China Sea, the total slope ranges from 0.2 to 0.5. But in the ocean areas west downwind from the southeast China continent, the slope presents in a higher range from 0.5 to 1.0, indicating a possible different emission ratio of NO_x to CO at Yangtze River Delta. Different from United States, the slope of O_3 to CO in China regions is almost all attributed to contribution from primary anthropogenic CO, while that of oxidation of biogenic isoprene is not significant. It is apparently that air mass with a homogeneous O_3 to CO slope distribution are coming from the southeast coast (coastal economic zones). These air masses are consequently transported to Japan and Korea and then affect the local troposphere O_3 variation in these countries.

On the continent of China, beside the coastal economic zones, primary anthropogenic CO also contributes to the slope of O_3 to CO at the northeastern China. However, low values display at the highest anthropogenic emissions regions: Beijing-Tianjin-Hebei economic zone, Yangtze River Delta and Pearl River Delta. Compared to primary anthropogenic CO, the contributions from oxidation of biogenic isoprene and anthropogenic VOCs are insignificant.

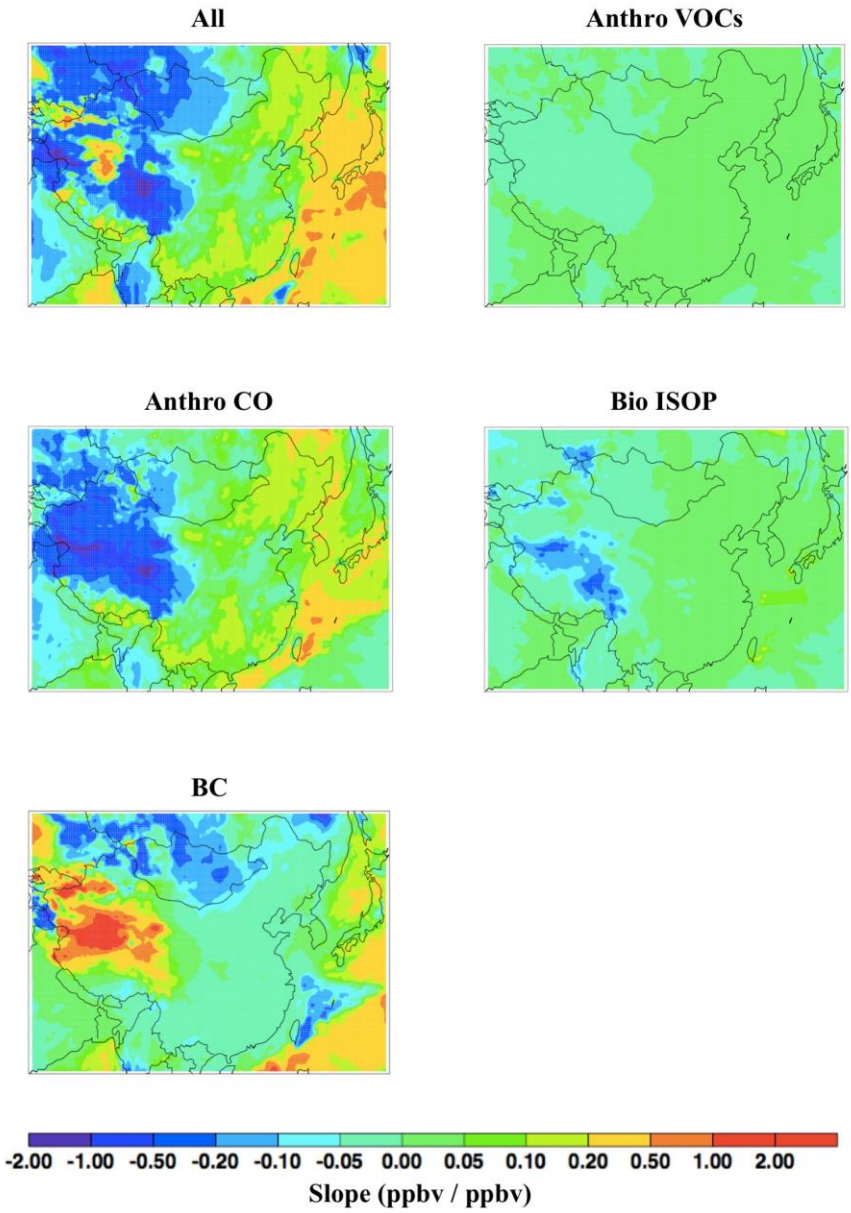

Figure 2.13 – The same as Figure 2.11 but for the distribution of O_3-CO slope and sub slopes in China region.

The contribution by CO transported from the (lateral) boundary to the O_3-CO regression slope, the same as that in United States, only contributed at the mountain areas like Qinghai-Tibet Plateau and other remote background areas.

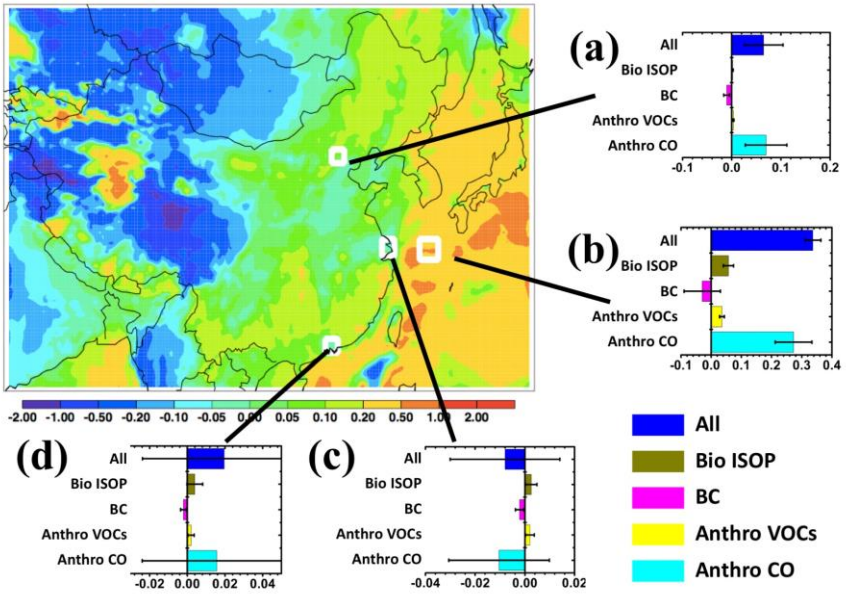

Figure 2.14 – Regional compositions of O_3-CO slope and sub slopes in China: (a) Shanghai, (b) East China Sea, (c) Pearl River Delta Economic Zone, (d) Beijing.

Regional O_3-CO slope and sub slopes' compositions in some typical areas in China are shown in Figure 2.14. The result indicates that no matter in north cities like Beijing (d), or the south cities like Shanghai (a) and Pearl River Delta Economic Zone (c), CO from primary anthropogenic emissions contribute the most. Correspondingly, the ocean regions (b) near Shanghai show that the O_3-CO slopes in air masses exported from the continent of China are controlled by primary anthropogenic emissions.

2.5 Summary and Implements

The extensive measurements in the lower atmosphere during the DISCOVER-AQ experiment over the Washington-Baltimore area in July 2011 are analyzed in this study to understand the contributions to the O_3-CO regression slope from different CO sources using REAM model. The observed vertical and temporal variations of O_3, CO, NO_x, isoprene, and CH_2O, as well as the correlation between O_3 and CO, and the O_3-CO regression slopes are reproduced well in the model. The observations show a robust O_3-CO regression in the boundary layer in daytime. We implement a new regression slope decomposition method with 4 simulated CO tracers, including primary anthropogenic emissions, oxidation of anthropogenic VOCs, oxidation of biogenic isoprene, and transport from model boundaries, to quantify their contributions to the O_3-CO regression slope.

We find that the O_3-CO regression slope in the boundary layer (300 m – 2.5 km) is controlled by primary anthropogenic CO emissions and CO production by isoprene oxidation. The contribution by the latter is nearly as large as the former. From 300 m to the surface, the model results indicate that the latter contribution is dominant. Previous studies discussed complicating factors in using the O_3-CO regression slope and the total anthropogenic CO emissions to estimate the net anthropogenic O_3 production amount [e.g., Chin et al., 1994; Parrish et al., 1998]. Due to the decrease of anthropogenic emissions, our study suggests that biogenic CO needs to be properly accounted for through modeling or measurement means for the eastern United States in the summer when using the O_3-CO regression slope method. The biogenic contribution will continue increasing as anthropogenic emissions decrease. At some point in the future, it can be expected that the biogenic contribution may become the most dominant. Despite the change of the relative

importance of anthropogenic and biogenic emissions, the consistent enhancement of O_3 relative to CO observed in the boundary layer, as indicated by the O_3-CO regression slope, provides a useful constraint on model photochemistry and emissions.

We also analyzed the factors that contribute to the regression slope of surface atmospheric O_3 and CO in United States and China regions. Biogenic isoprene oxidation makes a significant contribution to the slope of O_3 to CO in eastern United States. It provides nearly half of the contributions to the slope of O_3 to CO in the air mass exported from the eastern coast to the North Atlantic pollutant outflow region in summer time. This result revealed that it is not reasonable to use the slope of O_3 to CO with an anthropogenic CO emissions inventory to estimate the anthropogenic O_3 formation and transport from eastern United States to the North Atlantic and even to the global scale. At eastern United States of continent, the attribution of the total slope is primarily taken placed by CO from biogenic isoprene oxidation. However, slope of O_3 to CO in the air mass exported from the western coast of United States to the eastern Pacific are majorly contributed by anthropogenic emissions.

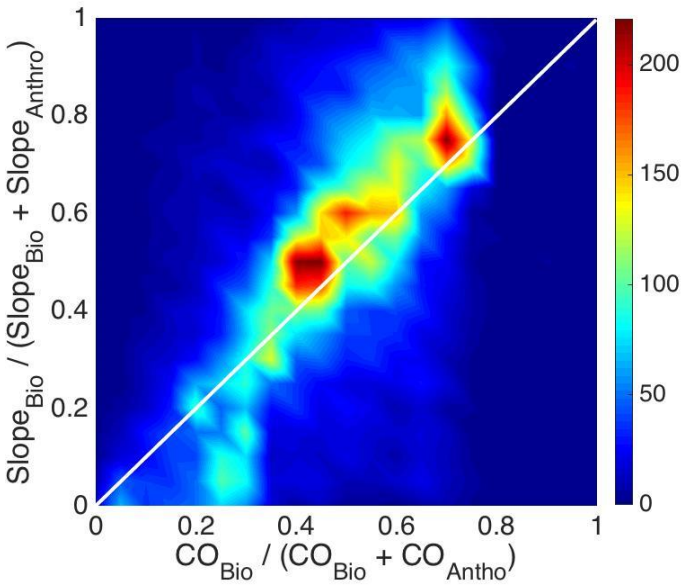

Figure 2.15 – Density plot of proportion of sub-slope of O_3 to CO biogenic isoprene oxidation to the sum of sub-slopes of O_3 to CO biogenic isoprene oxidation and primary anthropogenic CO, versus the proportion of CO from biogenic isoprene oxidation to the sum of CO from both of these two sources, where "Bio" and "anthro" donate "biogenic isoprene oxidation" and "primary anthropogenic emissions", respectively.

CO from biogenic isoprene oxidation and primary anthropogenic emissions are major two contributors to the slope of O_3 to CO in most areas. We compare the proportion of sub-slope of the former to the sum of sub-slopes of both, versus the proportion of CO from biogenic isoprene oxidation to the sum of CO from both of these two sources, by grouping result in each grid box (Figure 2.15). The contribution from biogenic isoprene oxidation to the slope of O_3 to CO is nearly linear positive correlated to its proportion of amount. This is the reason of the different sub-slope compositions between western and eastern of United States, where the biogenic isoprene emissions in the former is much less than the later [Millet et al., 2008; Guenther et al., 2012].

Anthropogenic emissions of CO decreased dramatically in the last decade (https://www.epa.gov/air-emissions-inventories/air-pollutant-emissions-trends-data). The large decrease of anthropogenic CO emissions can lead to the increase of proportion of CO from biogenic isoprene oxidation. We compare the absolute proportion of sub-slope of CO from biogenic isoprene oxidation and that of primary anthropogenic CO between 2005 and 2011 (Figure 2.16). The proportion of CO from biogenic isoprene oxidation increased ~75% in the past 6 years while that from anthropogenic CO decreased ~ 25%.

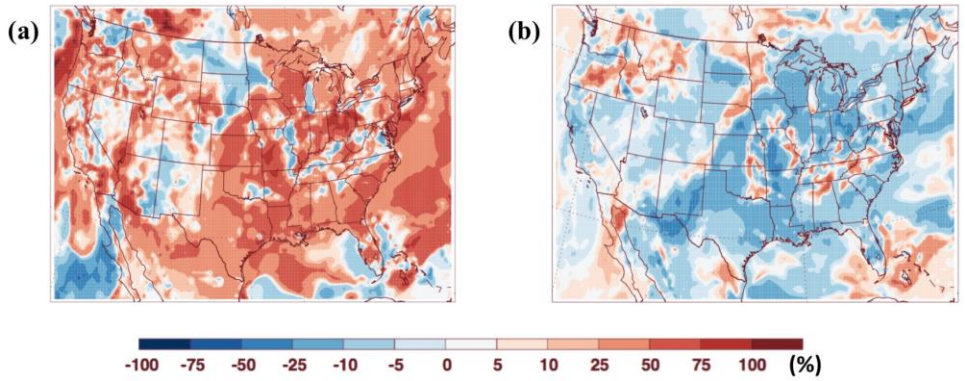

Figure 2.16 – Difference (2011 minus 2005) of absolute proportion of contribution to the slope of O_3 to CO from (a) biogenic isoprene oxidation and (b) primary anthropogenic emissions, to the sum of 4 contributions.

In China, since the anthropogenic emissions are still high as a result of the fast development of economy, and as well the vegetation coverage is lower compared with eastern of United States, the slope scenario is majorly anthropogenic CO emissions controlled in most land areas and oceans. However, with the emission reduction progress in the future, a larger biogenic contribution is expected.

CHAPTER 3. ESTIMATOR OF SURFACE OZONE USING FORMALDEHYDE AND CARBON MONOXIDE CONCENTRATIONS OVER THE EASTERN UNITED STATES IN SUMMER

3.1 Introduction

O_3 is a major pollutant in the troposphere [e.g., Logan et al., 1981; Wang and Jacob, 1998; Wang et al., 1998; Lelieveld and Dentener, 2000]. Thus, monitoring tropospheric O_3 at regional and global scales is important for environmental protection. Space-borne remote sensing utilizing its absorption features in the ultraviolet (UV) and the thermal infrared (TIR) bands is the most convenient way to provide O_3 spatial distributions around the globe. However, because of the molecular scattering of UV [Liu et al., 2010] and lack of contrast of TIR [Beer, 2006], the satellite observations for O_3 still show a limited sensitivity in the lowermost troposphere, especially near the surface, which is directly relevant to air quality [e.g., Cuesta et al, 2013].

3-D air quality models can provide information on the distribution of surface O_3, but the accuracy of surface O_3 simulations is limited by uncertainties in precursor emissions, atmospheric processes, and nonlinear photochemistry. Some model uncertainties can be mitigated through probabilistic approaches [e.g., Dabberdt et al., 2004; Delle Monache et al., 2006a, b; Vautard et al., 2009]. Statistical methods have also been applied to estimate O_3 distributions, including classification and regression trees, linear regression, and neural networks [e.g., Burrows et al., 1995; Van der Wal and Janssen, 2000; Perez and Reyes, 2006; Cobourn et al., 2007; Shad et al., 2009, Biancofiore et al., 2015]. The advantage of such statistical modeling is that it offers moderate to high accuracy at a moderate cost

[Zhang et al., 2012a, b]. However, the nature of statistical modeling often requires a suite of input variables and does not enable better understanding of chemical and physical processes [e.g., Guillas et al., 2008].

An alternative to numerical or statistical modeling is to make use of the observations of O_3 surrogates. The correlations between O_3 and other chemical species might be used effectively to diagnose O_3 chemical and physical processes [e.g., Parrish et al., 1993; Chin et al., 1994; Wang and Zeng, 2004; Koo et al., 2012]. CO is often observed and simulated to have a linear relationship with O_3 in the lower atmosphere [e.g., Parrish et al., 1993; Chin et al., 1994; Buhr et al., 1996; Parrish et al., 1998; Cardenas et al., 1998; Li et al., 2002; Cooper et al., 2002a, b; Honrath et al., 2004; Mao and Talbot, 2004; Huntrieser et al., 2005; Cheng et al., 2017]. Over the eastern United States, the observed slope of O_3 to CO at ~0.3 reflects contributions by CO from primary anthropogenic emissions and that from biogenic isoprene oxidation [Cheng et al., 2017] due in part to the production of both O_3 and CO from biogenic VOCs oxidation [Guenther et al., 1995; Atkinson and Avey, 1998; Pierce et al., 1998; Hudman et al., 2008; Pang et al., 2009; Choi et al., 2010; Geng et al., 2011; Lee et al., 2014; Zhang and Wang, 2016]. It implies that the observations of CO concentrations can be potentially applied to track the contributions to surface O_3 by anthropogenic and biogenic factors. In addition to EPA surface monitoring networks, satellite observations of lower tropospheric CO are more promising than O_3 over polluted regions since CO concentrations are usually higher in the boundary layer than the free troposphere and unlike O_3 it does not have high concentrations in the stratosphere, although both CO and O_3 have substantial free tropospheric columns relative to the boundary layer. CO can be detected by satellite TIR, near infrared (NIR) and joint TIR and NIR instruments, such as the NIR SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY (SCIAMACHY) [e.g., De Laat et al., 2012] and the NIR and

TIR Pollution in the Troposphere (MOPITT) [e.g., Emmons et al., 2004; Straume et al., 2005].

CH_2O is a principal intermediate species in the oxidation of atmospheric hydrocarbons [e.g., Wiedinmyer et al., 2001; Duane et al., 2002; Pang et al., 2009 and references therein]. It is also a major radical source leading to ozone production in the presence of NO_x [e.g., Liu et al., 2012b]. We will show that regional CH_2O is also correlated to O_3 in section 3. Since the sources of CO and CH_2O are often different, the two correlations can provide separate constraints on O_3 distributions. While regulatory monitoring of surface CH_2O is unavailable, CH_2O is detectable from space with good sensitivities in the boundary layer by measuring backscattered solar UV between 325 and 360 nm [Chance et al., 2000], including SCIAMACHY [Wittrock et al., 2006], Ozone Monitoring Instrument (OMI) [Kurosu et al., 2004; González Abad et al., 2015], Global Ozone Monitoring Experiment–2 (GOME-2) [De Smedt et al., 2012], and Ozone Mapping Profiler Suite (OMPS) [Li et al., 2015; González Abad et al., 2016].

In this study, we apply REAM model to quantitatively study factors contributing to the observed correlations and regression slopes of O_3 with CH_2O and CO and those of CO with CH_2O using the DISCOVER-AQ measurements. The analysis will show that it is feasible to use observed aircraft CO and CH_2O concentrations to improve estimates of the surface O_3 distribution over the eastern United States. The uncertainties of current satellite-derived near-surface CO and CH_2O are still too large [e.g., Gloudemans et al., 2005 ; Buchwitz et al., 2007; Zhang et al., 2016b] to be applied in this method. The EPA surface CO monitoring data are also unusable since CO concentrations are often below the reporting limit [e.g., Zeng and Wang, 2011]. We therefore use the model simulated data, which reasonably captures the observed aircraft concentrations and relationships of O_3, CO, and CH_2O during the DISCOVER-AQ campaign, to develop and evaluate a surface O_3 estimation model with surface CO and CH_2O as input parameters. In the future,

geostationary satellite instruments such as TEMPO [Chance et al., 2013] will greatly improve the quality of near surface measurements of O_3 precursors and therefore make it possible to use observed CH_2O in the O_3 estimator we develop here. We describe the 3-D chemical transport model, DISCOVER-AQ dataset, analysis methods, the O_3 estimation model, and validation method in section 3.2. Section 3.3 describes the analysis, modeling, and evaluation results. Discussion of implementing the O_3 estimator and conclusions are given in section 3.4.

3.2 Data and analysis methods

3.2.1 Observations and Simulations

The observation data use in this study were obtained from the NASA 2011 DISCOVER-AQ airborne campaign (http://www-air.larc.nasa.gov/missions/discover-aq/discover-aq.html). Sampling by the NASA P-3B aircraft was conducted from Washington's Beltway northeast to Baltimore and continuing on to the Delaware state line and occasionally over the Chesapeake Bay. 14 flights over 6 locations selected for aircraft spirals were carried out to measure the vertical structures of pollutants. 253 daytime vertical profiles from 300 m to 5 km were measured between June 27 and July 31. CO was measured by a diode laser spectrometer [Sachse et al., 1987]. O_3 was measured by the National Center for Atmospheric Research 4-channel chemiluminescence instrument [Brent et al., 2013]. The uncertainties of the measurements on these two species are 2% and 5%, respectively. CH_2O was measured by a difference frequency generation absorption spectrometer [Weibring et al., 2010]. For CH_2O levels above 1 ppb the total measurement uncertainty at the 1σ level was estimate to be around 5%, which folds in systematic and limits of detection uncertainties. To evaluate model simulations with the observations, we

identify the model profiles corresponding to the locations of aircraft spirals and the time of aircraft sampling. Corresponding model vertical profiles and observations are used in correlation analysis to evaluate model performance.

We use a 3-D REAM to represent the observations and then conduct further correlation analyses. The model domain covers the contiguous United States with a horizontal resolution of 36 × 36 km^2. The chemistry mechanism in REAM is the GEOS-Chem standard chemical mechanism (V9-02) [Bey et al., 2001] with updates of kinetics data (http://jpldataeval.jpl.nasa.gov). The anthropogenic emission inventory used in the model is the 2011 National Emission Inventory (NEI) (https://www.epa.gov/air-emissions-inventories/2011-national-emissions-inventory-nei-data). The biogenic isoprene emissions are the results of the MEGAN version 2.1 [Guenther et al., 2012]. Initial and boundary conditions for chemical tracers are taken from the GEOS-Chem (V9-02) 2° × 2.5° simulation results [Bey et al., 2001]. Meteorology fields are from the WRF model results assimilated using the CFSR products (http://cfs.ncep.noaa.gov/cfsr).

The previous study by Cheng et al. [2017] shows that REAM simulates well the observed vertical and temporal variations of O_3, CO, NO_x, isoprene, and CH_2O, as well as the correlation between O_3 and CO, and the O_3-CO regression slopes during the 2011 DISCOVER-AQ campaign. In this work, we therefore focus on analyzing the correlation and regression slope of O_3-CH_2O. We trace separately via tagged tracers three different CH_2O sources: primary anthropogenic emissions and the oxidation of anthropogenic VOCs ($CH_2O_{anthroVOCs}$), the oxidation of biogenic isoprene ($CH_2O_{bioISOP}$), and transport from model lateral and upper boundaries (CH_2O_{BC}), to analyze the contribution from each source to the observed O_3-CH_2O relationship. Other biogenic VOCs are not taken into account because isoprene provides the source for the vast majority biogenic CH_2O [e.g., Kesselmeier and Staudt, 1999; Lathière et al., 2006; Guenther et al., 2012; Sindelarova et al., 2014]. In tagged-tracer simulations, relevant species and radicals, such as O_3, NO_x, and

HO_x (OH and HO_2), are fixed using results archived from the standard simulation. The sum of the three individual tagged tracers is within 2% of the total CH_2O concentrations in the standard simulation for grid cells over the Washington-Baltimore region. We carry out minor scaling adjustments in post-processing, assuming that relative CH_2O attributions stay the same, to ensure that the sum of the CH_2O tracers is the same as the total CH_2O for each grid cell in the standard simulation. With simulated CH_2O attribution results, we can decompose the O_3-CH_2O regression slope into 3 sub-slopes of the corresponding CH_2O tracers (Eq. (2), derived in the Appendix),

$$\text{Least} - \text{Squares Regression Slope of } O_3\text{-}CH_2O = \frac{Cov(CH_2O_{anthroVOCs}, O_3)}{Var(CH_2O_{total})} + \frac{Cov(CH_2O_{BC}, O_3)}{Var(CH_2O_{total})} + \frac{Cov(CH_2O_{bioISOP}, O_3)}{Var(CH_2O_{total})}, \quad (2)$$

where *Cov* and *Var* denote covariance and variance, respectively. Eq. (2) shows that the contribution of each CH_2O tracer, i.e., the sub-slope values, to the O_3-CH_2O regression slope is proportional to its covariance with O_3. It is therefore possible to have both positive and negative slope contributions.

For the DISCOVER-AQ region during the Baltimore-Washington study, where the majority of CH_2O is biogenic, we can use the regression slope of O_3-CH_2O in the evaluation of model results using the observations. When extending the analysis using Eq. (2) to remote regions, the small variance of background CH_2O leads to abnormally large slopes, making it difficult to show the spatial distribution of the regression slope. We therefor use an "inversed slope" of $\Delta CH_2O/\Delta O_3$ (Eq. (2), see the Appendix) to illustrate the spatial distribution over the US since the variance of O_3 is a more stable denominator that that of CH_2O,

$$\text{Least} - \text{Squares Regression Slope of } CH_2O\text{-}O_3 = \frac{Cov(CH_2O_{anthroVOCs}, O_3)}{Var(O_3)} + \frac{Cov(CH_2O_{BC}, O_3)}{Var(O_3)} + \frac{Cov(CH_2O_{bioISOP}, O_3)}{Var(O_3)}. \quad (3)$$

3.2.2 Surface ozone distribution estimation

The total O_3 concentration is contributed by three major sources: anthropogenic O_3 production ($O_{3_{anthro}}$), biogenic O_3 production ($O_{3_{bioISOP}}$) and the transport from the lateral and upper model boundaries ($O_{3_{background}}$) (Eq. (4)). In the ozone estimation model, we estimate $O_{3_{anthro}}$ and $O_{3_{bioISOP}}$ using the regressions of O_3 with source-tagged CO and CH_2O. We will show in the next section that the correlations and regression slopes of O_3-CH_2O and CO-CH_2O are almost entirely due to biogenic isoprene over the eastern U.S., where CH_2O concentrations are mostly due to oxidation of biogenic isoprene. We make use of this finding and use CH_2O as a proxy for O_3 related to biogenic emissions. We decompose surface O_3 concentrations in Eq. (4) into 3 components related to regional anthropogenic emissions, biogenic emissions, and background (not related to the emissions within the estimation domain). We approximate the emission related components using CO (Eq. (5)). Recognizing that we would like to make use of observation-based CO and CH_2O concentrations, we replace CO from anthropogenic emissions (CO_{anthro}) with CO-$CO_{biogenic}$-$CO_{background}$, where $CO_{biogenic}$ is CO from biogenic isoprene oxidation that proceeds through CH_2O, and further compute $CO_{biogenic}$ as a function of CH_2O from biogenic isoprene oxidation ($CH_2O_{bioISOP}$) (Eq. (6)). Considering that most of CH_2O is biogenic over the region in the summer, we replace $CH_2O_{bioISOP}$ as CH_2O-$CH_2O_{background}$ (Eq. (7)).

$$[O_3] = [O_3]_{\text{anthro}} + [O_3]_{\text{bigenic}} + [O_3]_{\text{background}} \quad (4)$$

$$\approx \frac{\Delta[O_3]}{\Delta[CO]_{\text{anthro}}} \times [CO]_{\text{anthro}} + \frac{\Delta[O_3]}{\Delta[CO]_{\text{bioISOP}}} \times [CO]_{\text{bioISOP}} + [O_3]_{\text{background}} \quad (5)$$

$$= \frac{\Delta[O_3]}{\Delta[CO]_{\text{anthro}}} \times \left([CO]_{\text{total}} - \frac{\Delta[CO]_{\text{bioISOP}}}{\Delta[CH_2O]_{\text{bioISOP}}} \times [CH_2O]_{\text{bioISOP}} - [CO]_{\text{background}}\right) + \frac{\Delta[O_3]}{\Delta[CO]_{\text{bioISOP}}} \times \frac{\Delta[CO]_{\text{bioISOP}}}{\Delta[CH_2O]_{\text{bioISOP}}} \times [CH_2O]_{\text{bioISOP}} + [O_3]_{\text{background}} \quad (6)$$

$$\approx \frac{\Delta[O_3]}{\Delta[CO]_{\text{anthro}}} \times \left([CO]_{\text{total}} - \frac{\Delta[CO]_{\text{bioISOP}}}{\Delta[CH_2O]_{\text{bioISOP}}} \times ([CH_2O]_{\text{total}} - [CH_2O]_{\text{background}}) - [CO]_{\text{background}}\right) + \frac{\Delta[O_3]}{\Delta[CH_2O]_{\text{bioISOP}}} \times ([CH_2O]_{\text{total}} - [CH_2O]_{\text{background}}) + [O_3]_{\text{background}} \quad (7)$$

$$= A \times [CO]_{\text{total}} + B \times [CH_2O]_{\text{total}} + C \quad (8)$$

where $A = \frac{\Delta[O_3]}{\Delta[CO]_{\text{anthro}}}$, $B = -\frac{\Delta[O_3]}{\Delta[CO]_{\text{anthro}}} \times \frac{\Delta[CO]_{\text{bioISOP}}}{\Delta[CH_2O]_{\text{bioISOP}}} + \frac{\Delta[O_3]}{\Delta[CH_2O]_{\text{bioISOP}}}$, $C = -\frac{\Delta[O_3]}{\Delta[CO]_{\text{anthro}}} \times [CO]_{\text{background}} + \left(\frac{\Delta[O_3]}{\Delta[CO]_{\text{anthro}}} \times \frac{\Delta[CO]_{\text{bioISOP}}}{\Delta[CH_2O]_{\text{bioISOP}}} - \frac{\Delta[O_3]}{\Delta[CH_2O]_{\text{bioISOP}}}\right) \times [CH_2O]_{\text{background}} + [O_3]_{\text{background}}$

We will show this function works well for the eastern US in the summer. The study domain (to be shown in Fig. 4) is selected where > 90% of monthly mean surface CH_2O is biogenic. For July 2011, background values do not have a significant spatial dependence and we use constants at 60 ppbv for $CO_{\text{background}}$, 10 ppbv for $O_{3\,\text{background}}$, and 200 pptv for $CH_2O_{\text{background}}$. Over the domain of surface O_3 estimation, we compute the $\frac{\Delta[O_3]}{\Delta[CO]_{\text{anthro}}}$, $\frac{\Delta[O_3]}{\Delta[CO]_{\text{bioISOP}}}$, and $\frac{\Delta[CO]_{\text{bioISOP}}}{\Delta[CH_2O]_{\text{bioISOP}}}$ values using the least-squares regression slope formulations

of O_3-CO by Cheng et al. (2017) and O_3-CH_2O in Eq. (2). As we will discuss in section 4, the unique feature of this surface O_3 estimation is that it is based on the temporospatial stability of the predicting parameters of the regression slopes and background values. To demonstrate the feasibility of this O_3 estimator, for hourly estimation from 11:00 A.M. to 4:00 P.M., we group the corresponding model data in all grid cells in the same hour and compute the regional regression slope values. Therefore, the estimated O_3 spatial variation is due to those of CO and CH_2O only. In order to test the reliability, the above estimation method is validated through "leave-one-out cross validation" (LOOCV). The method uses a single sample from the original datasets as the validation data, and the remaining samples (excluding the selected validation data point) are used in the estimator. Each sample in the datasets is used once as validation data. We conduct the validation individually for each hour of a day from 11:00 A.M. to 4:00 P.M.. Of the 31 days from July 1^{st} to July 31^{th} (d_1, d_2... d_{31}), the day i (d_i) is selected as the validation data, and the remaining days (d_1, d_2... $d_{(i-1)}$, $d_{(i+2)}$... d_{31}) are used to compute the regression slopes and estimate the surface O_3 distribution of day i. We exclude the data before 11:00 A.M. or after 4 P.M. when the estimation has large biases because photochemical production of O_3, CO, and CH_2O is still slow and the correlations among the species are not photochemically driven.

3.3 Results

3.3.1 Correlation and Regression Slopes of O_3-CH_2O and CO-CH_2O

We compare simulated O_3-CH_2O correlations to the DISCOVER-AQ observations as a function of altitude or as a function of time in Figure 3.1. Simulated and observed correlation coefficient (R) values are in good agreement. Strong correlation between O_3 and CH_2O is found from 0.3 km to the top of the boundary layer (~2.5 km) with a narrow

range of R value (~0.75). Near the surface, simulated R value shows a slight decrease due to the increase of the contribution of surface primary emissions of CH_2O, which are not directly related to photochemical O_3 production. From the upper boundary layer to the free troposphere, R value shows a drastic drop and changes sign from positive to negative. The sign change reflects due to the rapid decrease of CH_2O with altitude in the lower free troposphere where O_3 increases with altitude (Figure 2.3). We also compare simulated and observed O_3-CH_2O correlation coefficient as a function of time of the day with data from 300 m to 2.5 km. This R value represents the spatial correlation in a given hour and is somewhat lower than the temporal correlation just discussed. The model is in good agreement with the observations except the underestimation at 9 am and overestimation at 5 pm. In the middle and upper boundary layer, the relative contribution by photochemical CH_2O production increases. Therefore, the concurrent photochemical production of O_3 and CH_2O is likely the major factor contributing to the observed positive correlation between O_3 and CH_2O in the boundary layer.

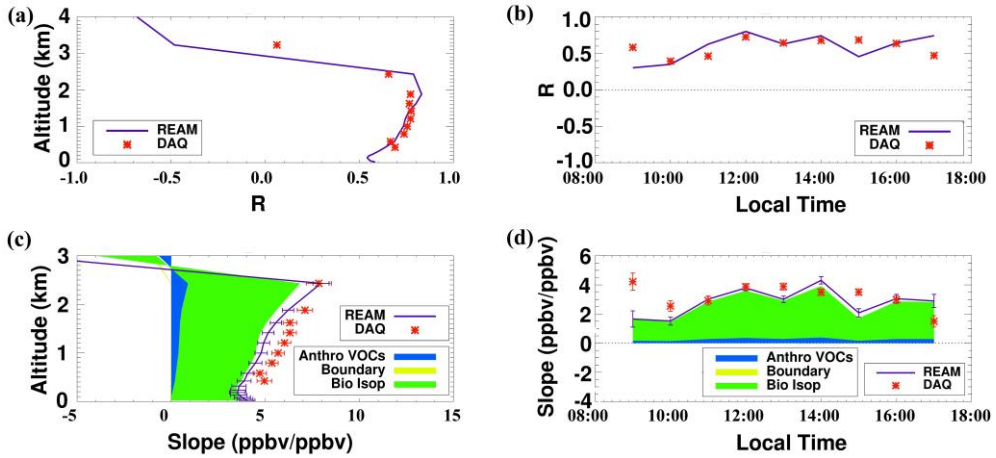

Figure 3.1 – Observed and simulated O_3-CH_2O correlation coefficients (R) and regression slope and sub-slopes (Eq. (2)) as a function of altitude (of data for 11 AM-4 PM, (a), (c)) and as a function of local time (of data for altitude of 0.3 - 2.5 km, (b), (d)). The sub-slopes due to varied CH_2O sources are shown using areas filled with different colors (Eq. (2)). The legends for different CH_2O sources are the same as in Eq. (2). The horizontal bars in (c) and vertical bars in (d) show the observed or simulated standard deviations of the regression slopes. The R, slope, and sub-slope values are computed using the DISCOVER-AQ observations or corresponding model data at a given altitude bin or for a given period.

The observed regression slope of O_3 to CH_2O is also captured by model simulation with satisfactory agreement in both vertical distribution and diurnal variation (shown in Figure 3.1 (c) and (d)). The O_3-CH_2O regression slope at a given altitude is underestimated by the model at 0.3-2.5 km by ~15%. However, the observed slight increase (from ~5 to ~8 ppbv/ppbv) in the boundary is captured by the model. The observed O_3-CH_2O regression slope at a given hour is better simulated than that at a given altitude. We further quantify source contributions by decomposing the O_3-CH_2O regression slope into the different CH_2O sources using Eq. (2). During DISCOVER-AQ, the regression slope of O_3-CH_2O is contributed almost exclusively by that of biogenic CH_2O.

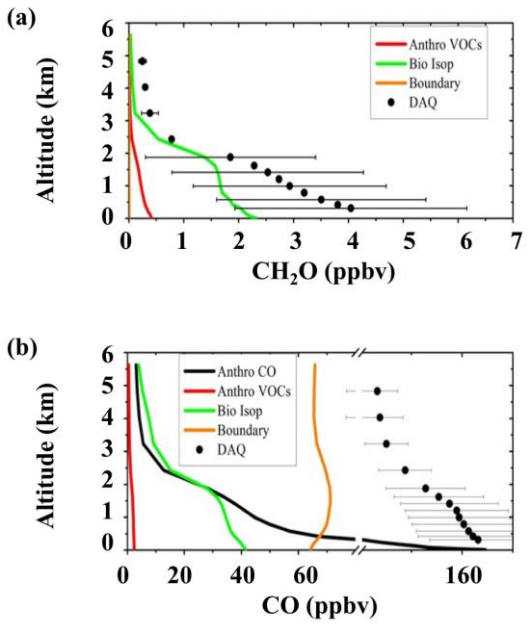

Figure 3.2 – Simulated and observed daytime (11 A.M. to 4 P.M.) vertical profiles of CH_2O and CO. "Anthro CO", "Anthro VOCs", "Bio ISOP", and "Boundary", denote CO due to primary anthropogenic emissions, anthropogenic VOC oxidation, biogenic isoprene oxidation, and transport from the model boundary, respectively. "DAQ" denotes DISCOVER-AQ observations.

Over the United States, we use the model results in Figure 3.3 to understand the relative contributions of CH_2O sources to the CH_2O-O_3 regression slope near the surface (Eq. (3)). The slope decomposition results show that the contribution from biogenic isoprene to the slope of O_3 to CH_2O is overwhelming over most regions of the eastern United States. Over the regions where it dominates, biogenic CH_2O has positive correlations with O_3 due to the concurrent production of these two species from the oxidation of biogenic VOCs. The exception is in central United States, where biogenic CH_2O concentrations are high but O_3 concentrations are low due to low NO_x concentrations. The chemical loss of O_3 leads to a negative regression slope, which seems high because of the low variance of O_3 in the region (Eq. (3). Without significant photochemical production,

the variation of O_3 in this region is low and is therefore relatively easy to estimate than the other regions. Overall, the short lifetimes of biogenic isoprene and CH_2O make it possible to use its concentrations to estimate the spatial variation of O_3 using relationships like Eq. (8).

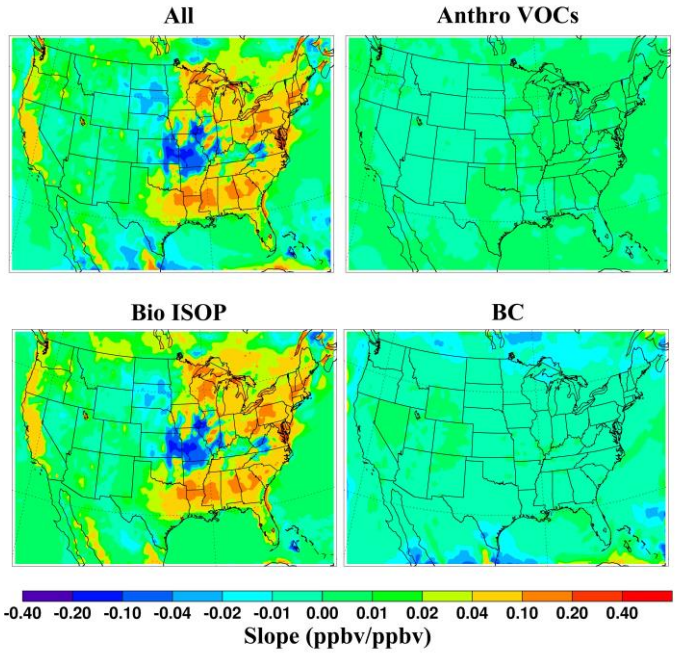

Figure 3.3 – Distribution of the regression slopes and sub slopes of surface CH_2O- O_3 (Eq. (3)) over the United States. The title "All" stands for the slope computed by total O_3 and total CH_2O concentration; the titles of "Bio ISOP", "Anthro VOCs", and "BC" stand for the sub slopes computed by total O_3 and CH_2O from primary anthropogenic emissions and oxidation of anthropogenic VOCs, oxidation of biogenic isoprene, and transport from the lateral and upper model boundaries, respectively. The slope and sub slope values on each grid are computed using grouped hourly data on daytime of 11: 00 A.M. – 04: 00 P.M. from July 1st to July 30th.

We also compare observed and simulated correlation and regression slopes of CO and CH_2O using an equation similar to Eq. (2) since it is used in Eq. (8). The simulation

results are in good agreement with the observations as a function of altitude or time of a day (Figure 3.4). Strong correlation between CO and CH_2O (R=~0.85) is observed and simulated from 0.3 km to 2.5 km (Figure 3.4 (a)). This correlation is the results of the co-emissions from anthropogenic sources and the co-production of mostly biogenic CO and CH_2O. In this altitude range, the regression slope of CO to CH_2O is about ~20 ppbv / ppbv without little variation (Figure 3.4 (c)). Below 0.3 km, simulated R value decreases towards the surface because the contribution of CO primary emissions increases significantly near the surface while the secondary formation is still the major sources of CH_2O (Figure 3.2). These two processes are not correlated. Correspondingly, the slope of CO to CH_2O increases from 0.2 km to the surface because the gradient of CO is larger than that of CH_2O (Figure 2.3) due to the surface primary emissions. However, we also test the correlation of surface $CO_{biogenic}$ and CH_2O who shows a R^2 in a range of 0.85 to 0.95 in the areas where more than 90% CH_2O is from biogenic isoprene oxidation. This reveals the reasonability of using CH_2O and the correlation of CO-CH_2O to separate anthropogenic and biogenic CO. Above 2.5 km, the R value decreases from the boundary layer top to the free troposphere. This is because above the boundary layer CO is mainly from transport from lateral and upper boundaries, which does not contribute to CH_2O as much due to its short chemical lifetime (Figure 3.2). The slope of CO to CH_2O also increases from the boundary layer to the free troposphere due to low concentrations of CH_2O in free troposphere. As a function of time in a day, the R value does not show a significant variation. The regression slope of CO-CH_2O remains at ~20 ppbv / ppbv in daytime except higher values in the morning (before 08: 00 A.M.) when solar radiation is weak.

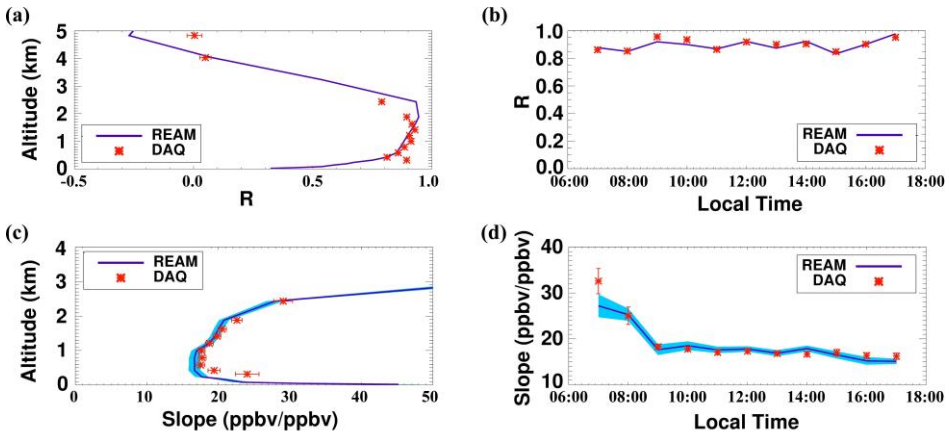

Figure 3.4 – Observed and simulated CO-CH_2O correlation coefficients (R) and regression slope for the Baltimore-Washington area as a function of altitude for daytime of 11: 00 A.M. – 04: 00 P.M. ((a), (c)) and local time for altitude of 0.3 - 2.5 km ((b), (d)). The horizontal bars in (c) and vertical bars in (d) show the observed standard deviations of the regression slopes. Shaded blue areas in (c) and (d) show simulated standard deviations of the regression slopes. The R and slope values are computed using the DISCOVER-AQ observations or corresponding model data at a given altitude bin or for a given time period.

3.3.2 Surface Ozone Estimation Using Eq. (8)

For July, 2011, we estimate the LOOCV surface O_3 distribution using Eq. (8). To examine the sensitivity of the estimation to emissions, we also increase (or decrease) anthropogenic emissions by 15%, 30%, and 50% respectively. The full REAM model is run with different emissions. For the LOOCV estimation using Eq. (8), the $\frac{\Delta[\mathbf{O_3}]}{\Delta[\mathbf{CO}]_{\mathbf{anthro}}}$, $\frac{\Delta[\mathbf{O_3}]}{\Delta[\mathbf{CO}]_{\mathbf{bioISOP}}}$, and $\frac{\Delta[\mathbf{CO}]_{\mathbf{bioISOP}}}{\Delta[\mathbf{CH_2O}]_{\mathbf{bioISOP}}}$ values are the same as in the standard simulation. The only changes are for the CO and CH_2O surface concentrations used. The premise is that the regression slopes of $\frac{\Delta[\mathbf{O_3}]}{\Delta[\mathbf{CO}]_{\mathbf{anthro}}}$, $\frac{\Delta[\mathbf{O_3}]}{\Delta[\mathbf{CO}]_{\mathbf{bioISOP}}}$, and $\frac{\Delta[\mathbf{CO}]_{\mathbf{bioISOP}}}{\Delta[\mathbf{CH_2O}]_{\mathbf{bioISOP}}}$ are relatively stable with respect to emission changes and therefore the estimation model using Eq. (8) provide a

robust means to estimate surface O_3 distribution when the distributions of CO and CH_2O are known.

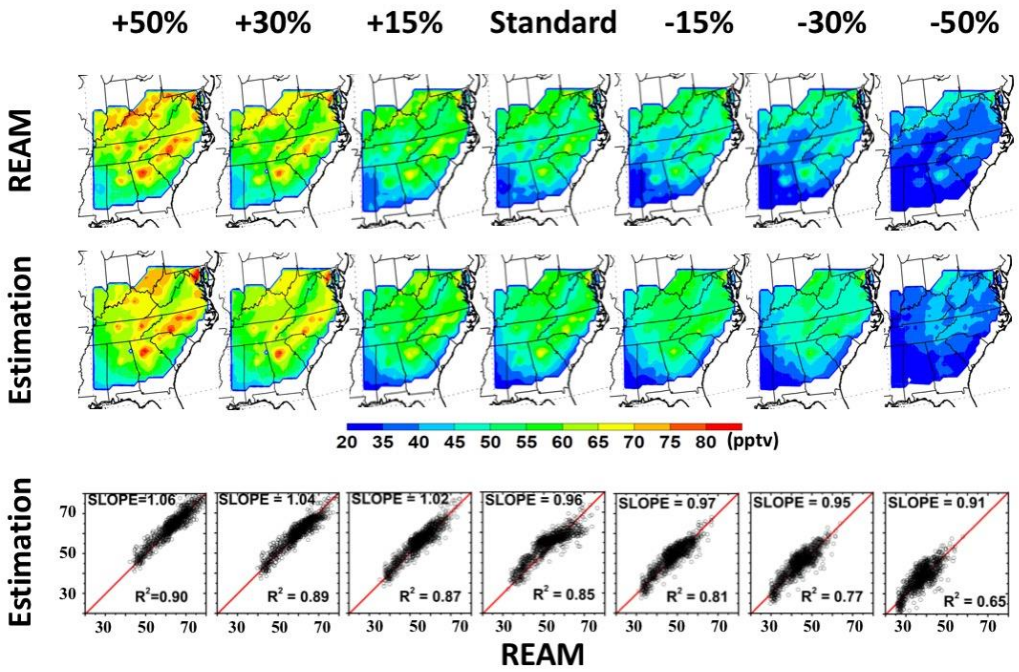

Figure 3.5 – Distributions of monthly mean (11:00 A.M. to 4:00 P.M.) REAM and LOOCV estimation of surface O3 concentrations for July,, 2011 under different anthropogenic emission scenarios. In LOOVC hourly estimation, the regional parameters in Eq. (8) are estimated using data not including the day of estimation. Scatterplots of corresponding grid-cell hourly REAM and LOOCV estimation data are shown in the third row; the 1:1 line is shown in red. 7 emission scenarios are presented. "+50%", "+30%", "+15%", "Standard", "-15%", "-30%", "-50%" on the top of columns denote 150%, 130%, 115%, original, 85%, 70% and 50% of anthropogenic CO and NOx emissions of NEI 2011.

Since Eqs. (4)-(7) are more accurate when CH_2O is dominated by oxidation of biogenic isoprene, the estimation evaluation is only for the eastern US regions where monthly mean biogenic CH_2O is > 90%. We compare the averaged estimation results with REAM results for these regions under different emission scenarios in Figure 3.4. In the standard simulation, the estimation shows a similar distribution and explains 85% of the

variance of the full REAM results (R^2 = 0.85) with no significant bias (Table 3.1). As the anthropogenic emissions decrease by 15-50%, the REAM model shows a decrease from 3.0 to 12.4 ppbv on average in these regions. The estimation model using Eq. (8) overestimates the O_3 decrease by 1.0-1.5 ppbv on average and the explained variance decreases from 81% to 65%. When anthropogenic emissions increase by 15%-50%, the full REAM results show surface O_3 increases from 2.7 to 11.8 ppbv. As anthropogenic emissions increase, the O_3 "hotspots" due to urban emissions become more obvious. The estimation model using Eq. (8) shows similar features but overestimates the O_3 increase by 0.7-1.1 ppbv on average. However, the explained variance (R^2 value) increases from 0.87 to 0.90 due largely to good estimations of urban increases. We also tested the estimation model performance by increasing (or decreasing) biogenic isoprene emissions by 15%, 30%, and 50%, respectively. The full REAM model shows much lower dependence of surface O_3 to biogenic emissions than anthropogenic emissions (Figure 3.5). The estimation model shows similar results with R^2 values ranging from 0.77 to 0.86 (Table 3.2).

Table 3.1 – Comparison of hourly REAM and LOOCV estimation of surface O_3 concentrations under different anthropogenic emissions scenarios. All grids in the domain in all certain hours (11:00 A.M. to 4:00 P.M.) are grouped. "Std" donates the standard deviation. "MSE" denotes the mean squared error.

Emission Scenarios	REAM Mean ± Std (ppbv)	Estimation Mean ± Std (ppbv)	Mean Bias	MSE	R^2	Least-squares Slope (ppbv/ppbv)
+50%	63.0 ± 8.2	64.1 ± 8.1	1.1	7.6	0.90	1.06
+30%	59.0 ± 7.8	59.9 ± 7.6	0.9	7.1	0.89	1.04

+15%	53.9 ± 7.5	54.6 ± 7.5	0.7	8.5	0.87	1.02
Standard	51.2 ± 6.7	51.8 ± 6.9	0.6	9.2	0.85	0.96
-15%	48.2 ± 6.8	47.2 ± 6.8	-1.0	9.5	0.81	0.97
-30%	44.3 ± 6.6	43.2 ± 6.3	-1.1	11.0	0.77	0.95
-50%	38.8 ± 6.2	37.3 ± 5.5	-1.5	13.6	0.65	0.91

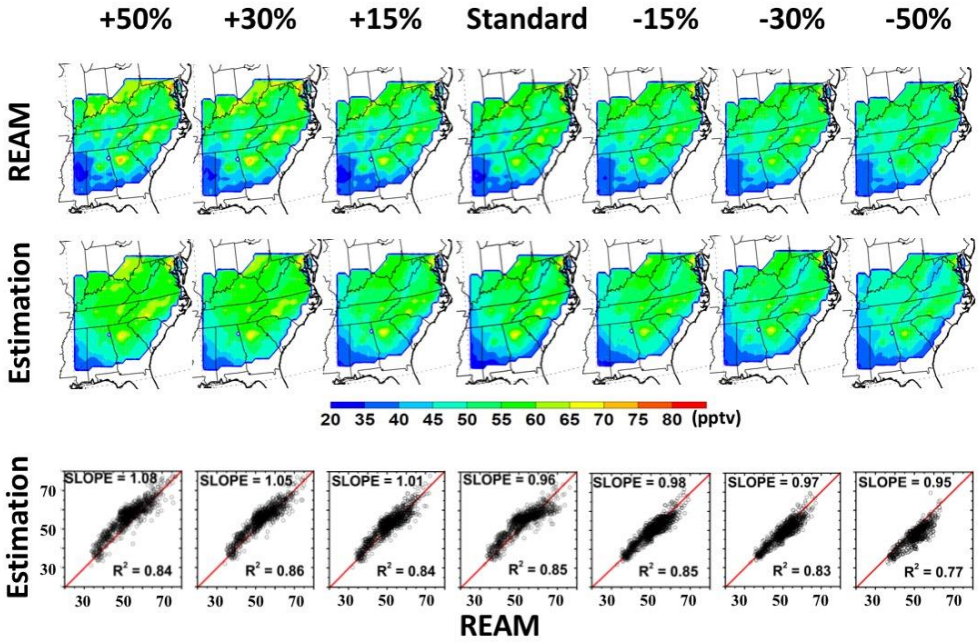

Figure 3.6 – Same as Figure 3.5 but for REAM and LOOCV estimations using Eq. (8) with 150%, 130%, 115%, original, 85%, 70% and 50% of the biogenic isoprene emissions in July 2011. The standard anthropogenic emissions are used.

Table 3.2 – Comparison of hourly (11:00 A.M. to 04:00 P.M.) simulated and predicted LOOCV surface O_3 in different biogenic isoprene emissions scenarios.

Emission Scenarios	REAM Mean ± Std (ppbv)	Estimation Mean ± Std (ppbv)	Mean Bias	MSE	R^2	Least-squares Slope (ppbv/ppbv)
+50%	51.6 ± 8.4	54.1 ± 8.1	2.5	17.7	0.84	1.08
+30%	52.4 ± 8.0	53.4 ± 7.5	1.0	12.3	0.86	1.05
+15%	50.3 ± 7.5	50.9 ± 7.1	0.6	10.6	0.84	1.01
Standard	51.2 ± 6.7	51.8 ± 6.9	0.6	9.2	0.85	0.96
-15%	50.9 ± 6.6	50.3 ± 6.1	-0.6	7.7	0.85	0.98
-30%	50.8 ± 6.2	49.0 ± 5.8	-1.7	9.2	0.83	0.97
-50%	50.0 ± 5.6	47.8 ± 5.3	-2.2	13.0	0.77	0.95

3.4 Summary and Implements

The surface O_3 estimator (Eq. (8)) works very well for the regions shown in Figure 3.4 and it is also quite robust with R^2 values > 0.77 for anthropogenic emissions in the range of -30% to +50% and biogenic emissions of -50% to +50%. Therefore, it has the potential of being used for rapid O_3 distribution assessment if the distributions of surface CO and CH_2O are known. Given the current lack of usable surface CO and CH_2O observations, we cannot test the estimator using observed data. In this analysis, the key parameters of the estimator, i.e., the regression slopes of

$\frac{\Delta[\mathbf{O_3}]}{\Delta[\mathbf{CO}]_{\mathbf{anthro}}}$, $\frac{\Delta[\mathbf{O_3}]}{\Delta[\mathbf{CO}]_{\mathbf{bioISOP}}}$, and $\frac{\Delta[\mathbf{CO}]_{\mathbf{bioISOP}}}{\Delta[\mathbf{CH_2O}]_{\mathbf{bioISOP}}}$ are based on model simulations although we show that model simulations are in good agreement with DISCOVER-AQ observations [Figs. 1 and 3; Cheng et al., 2017]. If the distributions of O_3, CO, and CH_2O are known, Eq. (8) can be used to obtain the observation-based regression slopes using least-squares regression. These parameters can provide insights in understanding of biases of air quality model simulations and be applied to improve the model.

Surface CO measurements are readily obtained if existing instruments are calibrated and the reporting limit is lowered. In situ observation of CH_2O is more complex and expensive than CO. For both species, the high concentrations in the boundary layer and very low concentrations in the stratosphere imply that satellite instruments have better sensitivity to derive their near-surface concentrations than O_3. The relatively large uncertainties of the current generation instruments on sun-synchronous satellites can be greatly reduced (due in part to the large increase of observation frequency) by instruments onboard geostationary satellites such as TEMPO over North America [Chance et al., 2013], SENTINEL-4 over Europe [Ingmann et al., 2012], GEMS over East Asia [Bak et al., 2013]. High-quality CH_2O measurements over the United States will be available from TEMPO. Deployment of near IR and thermal instruments on geostationary satellites will be needed for improved satellite measurements of near-surface CO. More complex statistical methods can be applied to combine such derived O_3 estimation with in situ surface O_3 observations.

There are limitations in the estimator of Eq. (8). The relatively good performance of the estimation model with changing anthropogenic emissions is not because surface O_3 is insensitive to anthropogenic NO_x emissions. In fact, most of the changes shown in Table

1 are due to NO_x. In REAM simulations, we make an implicit assumption that NO_x and CO emission ratios of anthropogenic sources do not change. When this assumption in invalid, for example, emission changes of specific sources that only affect either NO_x or CO, the estimator needs to be reconstructed with updated regression slopes of $\frac{\Delta[\mathbf{O_3}]}{\Delta[\mathbf{CO}]_{\mathbf{anthro}}}$, $\frac{\Delta[\mathbf{O_3}]}{\Delta[\mathbf{CO}]_{\mathbf{bioISOP}}}$, and $\frac{\Delta[\mathbf{CO}]_{\mathbf{bioISOP}}}{\Delta[\mathbf{CH_2O}]_{\mathbf{bioISOP}}}$. The estimator is valid only in the regions where CH_2O is dominated by oxidation of biogenic isoprene. As anthropogenic emissions are expected to continue decreasing [Cheng et al., 2017], the regions where the estimator can be applied will increase.

To summarize, in this study, extensive measurement of O_3, CO, and CH_2O were conducted during the DISCOVER-AQ aircraft experiment in July 2011 over the Washington-Baltimore area. We find strong correlation and stable linear regression slopes of O_3-CH_2O, O_3-CO and CO-CH_2O with no significant variation with time or altitude in the boundary layer. The concentrations, correlations and regression slopes of these tracers are reproduced well by the REAM model. We find that biogenic isoprene oxidation makes most of the contribution to the regression slopes of CH_2O-O_3 in large regions of the eastern United States using the slope decomposition method by tracing separately three different CH_2O sources, including primary anthropogenic emissions and oxidation of anthropogenic VOCs, oxidation of biogenic isoprene, and transport from the lateral and upper model boundaries.

Making use of the of the robust regression slopes, we construct a surface ozone estimation model using the distributions of CH_2O and CO as input parameters. In this model，CH_2O is used as a proxy to calculate O_3 and CO produced by the oxidation of

biogenic VOCs. The estimator can explain >77% of the surface O_3 variance simulated by the full 3-D model in the range of 70% to 150% of the anthropogenic emissions. It provides a fast regional surface O_3 estimation in most regions of the eastern United States in summer where CH_2O is dominated by oxidation of biogenic isoprene. With high-quality geostationary satellite observations of CO and CH_2O, the estimator could be applied to improve surface O_3 distribution, which is challenging to measure directly from space. The function of Eq. (7) relates surface concentrations of O_3 to those of CO and CH_2O. Using a sizable observation dataset of these concentrations of a given region, 6 regional metrics, including the regression slopes of $\frac{\Delta[O_3]}{\Delta[CO]_{anthro}}$, $\frac{\Delta[O_3]}{\Delta[CO]_{bioISOP}}$, and $\frac{\Delta[CO]_{bioISOP}}{\Delta[CH_2O]_{bioISOP}}$, and regional background concentrations of O_3, CO, and CH_2O, can be empirically determined and applied to investigate model performance and biases. Therefor the estimation model provides the means of using observations to evaluate and improve full-fledged 3-D air quality models.

CHAPTER 4. LARGE ALKANES EMISSIONS UNDERESTIMATIONS AND THEIR EFFECTS ON SUMMER OZONE AND FORMALDEHYDE MODELING IN HOUSTON AND DENVER

4.1 Introduction

The Houston metropolis is one of the most industrialized areas of the United States, as it is the home of a massive petrochemical complex in the Port of Houston and other sites near the Houston Ship Channel (HSC). Volatile organic compounds (VOCs) inventories of industry in this region are of concern because of their toxicity and the role they play in the formation of secondary air pollutants, such as secondary organic aerosols and tropospheric O_3 [e.g. Olaguer et al., 2009]. However, anthropogenic VOCs emissions due to transmit, loading and unloading, storage and domestic transportation, a potential problem in every petrochemical industry including any product distribution sites, are currently not addressed neither by the directive on the integrated pollution prevention or by other environmental regulations [Milazzo et al., 2017]. Their poorly quantified in emission inventories fail to quantify continues industrial VOCs emissions and that reliable estimates of these emissions can currently only be obtained from measurements. Severely underestimated VOCs emissions in the Houston areas, often by up to an order of magnitude, has been found in many researchs [Kleinman et al., 2002; Karl et al., 2003; Ryerson et al., 2003; Wert et al., 2003; Jobson et al., 2004; De Gouw et al., 2009; Parrish et al., 2009; Mellqvist et al., 2010; Washenfelder et al., 2010]. Johansson et al [2013] found alkanes, as high as 90% of VOCs emitted from oil operations, are underestimated by a factor of 5-15, due to the missing sources from oil and gas industries in the HSC area.

In Houston area, this inventory inaccuracy can finally result in underestimation of the burden of intermediates in the air (we will show this in Figure 4.6 and Figure 4.7), for example, CH_2O, an important tracer in the oxidation of atmospheric hydrocarbons in O_3 photo-chemical processes [Wiedinmyer et al., 2001; Duane et al., 2002; Pang et al., 2009]. It is considered as the major source for hydroperoxyl radical (OH) in the troposphere leading to O_3 production in the presence of NO_x. Besides direct emissions from combustion processes such as biomass burning, industrial flares and motor vehicles [Olaguer et al., 2009], CH_2O also originates from the decomposition of both biogenic and anthropogenic VOCs. Thus, a comparable simulation result of CH_2O from a model to the observations is an important examine of regional emissions estimation.

Several researchers have previously applied source apportionment techniques to VOCs data to study source receptor relationships in the Port of Houston [Kim et al., 2003; Buzcu and Fraser, 2006; Xie and Berkowitz, 2006; Leuchner and Rappenglück 2009; Guven Olaguer, 2010]. However, formaldehyde was not included in these analyses. Previous studies examined statistical correlations between CH_2O and CO, and CH_2O and O_3 [Friedfeld et al., 2002], or performed multiple regression analyses of CH_2O, peroxyacyl nitrates, sulfur dioxide and CO [Rappenglück et al., 2010], and indicated that secondary CH_2O is more significant compared to primary CH_2O generated by VOCs decomposition in the Houston air shed. An earlier modeling study [Luecken et al., 2006] determined that over 80% of CH_2O was due to secondary production.

The state of Colorado also has had a long history of natural resources extraction. The Denver-Julesburg Basin in Northeast Colorado produces both oil and natural gas (O&NG) from mostly tight sand and shale formations and has experienced rapid development of oil and gas drilling operations in recent years. The industrial equipment required for O&NG operations includes diesel trucks, drilling rigs, power generators, phase separators, dehydrators, storage tanks, compressors, and pipelines. Each piece of

equipment used to install, operate, or service a well is a known or potential emission source of VOCs. Gilman et al. [2013], through statistical regression analysis showed that more than half of the VOC-OH reactivity was attributable to emissions from O&NG operations indicating that these emissions are a significant source of O_3 precursors. However, current VOCs emissions inventory of Colorado also shows significant underestimation. For example, Pétron et, al. [2014] found the state inventory for total VOCs emitted by O&NG activities is at least a factor of 2 lower. Similar with Houston area, enhanced levels of C_2−C_5 alkanes have been observed in ambient air samples collected near areas of O&NG in Northeast Colorado and were attributed to primary emissions from the oil and gas industry [Gilman et al., 2013].

The primary objectives of this study are to propose corrections on the underestimations of VOCs emissions of current inventory in HSC and Denver, and to examine the corrections by comparing tracers like O_3, NO_x, CH_2O and CO, using a 3-D chemistry and transport model with observations. We describe the model, observational dataset, and corrections schemes in section 4.2. Simulation and comparison results of VOCs and other tracers are described in section 4.3. Conclusions and implications are given in section 4.4.

4.2 Data and Methods

4.2.1 Model and observational data

For the regional study of urban scale on Houston and its ship-channel, we use a REAM model with a higher horizontal resolution of 4×4 km^2. Transport is driven by the WRF model with the same horizontal resolution of 4×4 km^2, assimilated meteorological fields constrained by the CFSR products (http://cfs.ncep.noaa.gov/cfsr). The WRF model

domain is 10 grid-cell larger than that of REAM on each lateral side. The chemistry mechanism in REAM is adopted from the GEOS-Chem model (V9-02) [Bey et al., 2001] with updates of kinetics data (http://jpldataeval.jpl.nasa.gov). The anthropogenic emission inventory used in the model is the 2011 NEI (https://www.epa.gov/air-emissions-inventories/2011-national-emissions-inventory-nei-data). The biogenic isoprene emissions are computed using the MEGAN version 2.1 [Guenther et al., 2012]. Initial and boundary conditions for chemical tracers are taken from the REAM results of horizontal resolution of 36×36 km^2, whose setups are the same with those in Captor 2, and 3, but for the period of this study of interest.

The observational tracers' data (O_3, CO, NO_x, CH_2O and isoprene) for Houston were obtained during the NASA 2013 DISCOVER-AQ airborne campaign, sampling at Houston metropolitan area ranging from Conroe in the north to Galveston in the south. The vertical structures of pollutants were measured through 195 daytime vertical profiles from 300 m to 4.6 km over six selected locations during 13 flights by NASA's P-3B aircraft between August 28th and September 29th. CO was measured by a diode laser spectrometer [Sachse et al., 1987] with a 2% uncertainty. NO, NO_2, and O_3 were measured by the National Center for Atmospheric Research 4-channel chemiluminescence instrument [Brent et al., 2013] with 10%, 15% and 5% uncertainties, respectively. CH_2O was measured by a difference frequency generation absorption spectrometer [Weibring et al., 2010] with a 4% uncertainty. Isoprene was measured by a proton-transfer-reaction mass spectrometer [Lindinger et al., 1998] with a 10% uncertainty. There are 195 aircraft spirals around 8 locations. To evaluate model simulations with the observations, we first identify the model grid cells corresponding to the locations of aircraft spirals. These model profiles are then archived at the time of aircraft sampling. The corresponding aircraft and model data are used in correlation analysis and the averaged model vertical profiles are also compared to the observations.

Meteorology conditions are important for an urban scale study. Temperatures, and winds speed and directions in Houston were measured as well, and compared with WRF results (Figure 4.1 (a) (b)). WRF well captures the temperatures vertical profiles from 300 m to 5 km, except the underestimation from below 300m. Winds directions and speeds simulated by WRF are in good agreement with the observation.

The VOCs observational data for Houston were obtained through the Automated Gas Chromatographs (AutoGCs) system in Texas Commission on Environmental Quality (TCEQ). We identify the model grid cells corresponding to the locations on observation sites. Averaged diurnal variance of model results and the corresponding observational data are compared. Details of the TCEQ sites locations could be found at (https://www.tceq.texas.gov/airquality/monops/sites/air-mon-sites).

The observational tracers' data (O_3, CO, NO_x, CH_2O and isoprene) for Colorado were obtained during the NASA 2014 DISCOVER-AQ) airborne campaign from July 16th to August 16th. Sampling was token by NASA P-3B along the Northern Front Range from the Denver metropolitan area in the south to Fort Collins in the north extending eastward from the mountains as far as Greeley. There are 220 aircraft spirals around 8 locations (Figure 4.6). Vertical distribution of gaseous and particulate pollution was profiled from 300 m to 4.6 km above the surface over selected monitoring sites on the ground. The monitoring methods used were same with those in 2013 DISCOVER-AQ airborne campaign in Houston. The VOCs observational data is obtained by collaborating with a second study, the Front Range Air Pollution and Photochemistry Experiment (FRAPPE) which includes the C-130 research aircraft, as well as additional activities on the ground.

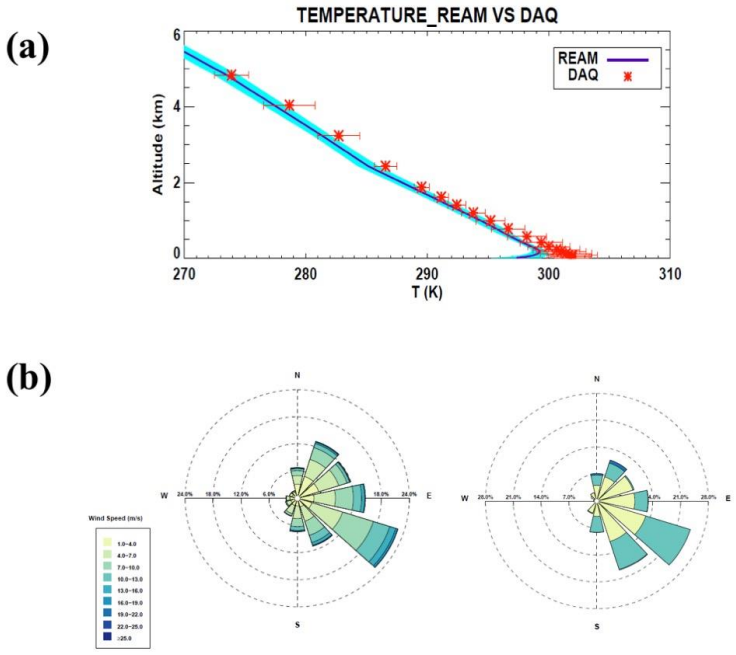

Figure 4.1 – Comparisons of (a) temperatures of WRF results and DISCOVER-AQ data, and (b) wind rose graphs of WRF results and DISCOVER-AQ data. In (a), DAQ donates DISCOVER-AQ data. In (b), each dashed circle presents the frequency of the wind with the corresponding direction range during the study time period.

4.2.2 VOCs Emissions Correction Scheme for Houston

In Figure 4.7 we will show that alkanes simulated by REAM in Houston are 70% to 90% lower than the TCEQ measurements based on current emissions inventory of NEI 2011. And in Figure 4.7 we will show that CH_2O is simulated 10% to 40% lower than the observational data. Both two gaps indicate the existence of a large VOCs missing. To test whether the VOCs emissions of transmit, loading and unloading, storage and domestic transportation are the missing sources, we figure out all locations of the obvious shipping ports from satellite image (Figure 4.2). Model grid cells which cover these ports and corresponding storage facilities are selected (Figure 4.2) to conduct the emissions corrections. Since alkanes has the majority proportion in VOCs emissions from oil and gas

operations and several research found that higher alkanes (C>4) that are important CH_2O precursors [e.g. Johansson et al, 2013]. We simplify the correction of emissions by choose a certain factor to be multiplied to the alkanes emissions on these selected grid cells. Averaged model results with corrected emissions will compared to that from original emissions and to observational data.

Two schemes of emissions correction supposed by recent studies are compared: (1) a factor of 5 is used to correct alkanes emissions at the whole Houston metropolitan area, according to the correction proposed by Zhu et al. [2014]. (2) a factor of 15 is used at the selected grid cells (Figure 4.2), according to Johansson et al. [2014]. Simulated and observed higher alkanes (C>4) are compared at 3 different ranges: (1) sites within 5 km to the center of Houston urban, (2) sites ranges of 5 km to 15 km from the center of Houston urban, and sites further than 15 km from the center of Houston urban.

Simulation results using the scheme (2), with 15 as the factor multiplied at selected grid cells, shows a better agreement with the observation data. Compared to the other one scheme multiplying a factor of 5 at the whole Houston metropolitan area, Butane (Figure 4.3), pentane (Figure 4.4) and larger alkanes (C>5) (Figure 4.5) present obvious increases of 50% at the sites further than 15 km from the center of Houston urban area and decreases of 20% to 70% at those within 15 km. The major difference is at night time where VOCs cumulate in the absence of photochemistry. The simulated results indicated that scheme (2) can better capture the characters of the distribution of alkanes at the area of interest. In the following simulations, schemes (2) is used to correct the emissions of NEI 2011.

Figure 4.2 – Locations of ship/vessel loading ports (yellow "thumbtack") with adjoining extensive storage and handling facilities, and the model domain of REAM (red dashed rectangles) which cover the former. The distribution (yellow dashed circles) shows the range of <5 km, 5km-10km and >15km to the Houston urban frontier (end of the ship channel).

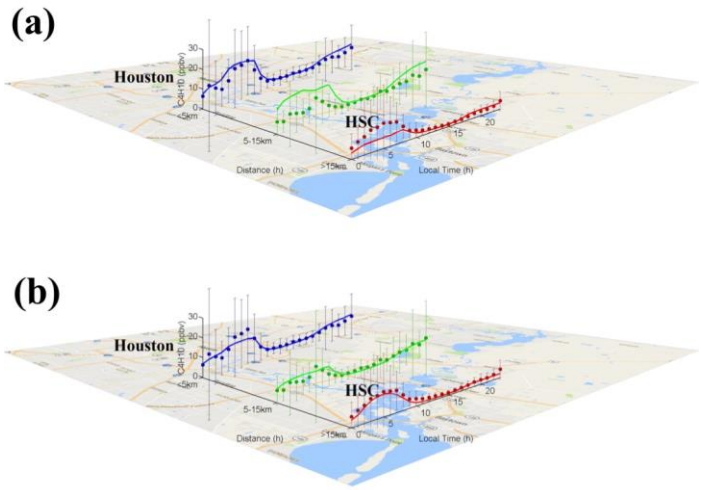

Figure 4.3 – Observed and simulated hourly mixing ratios of butane for TCEQ sites with distances of less than 5km, 5km-15km and larger than 15km to the west end of ship channel. Evenly correction of emission inventory of NEI 2011 (a) and correction

only at shipping ports further (b) are shown. The horizontal bars show the observed standard deviations.

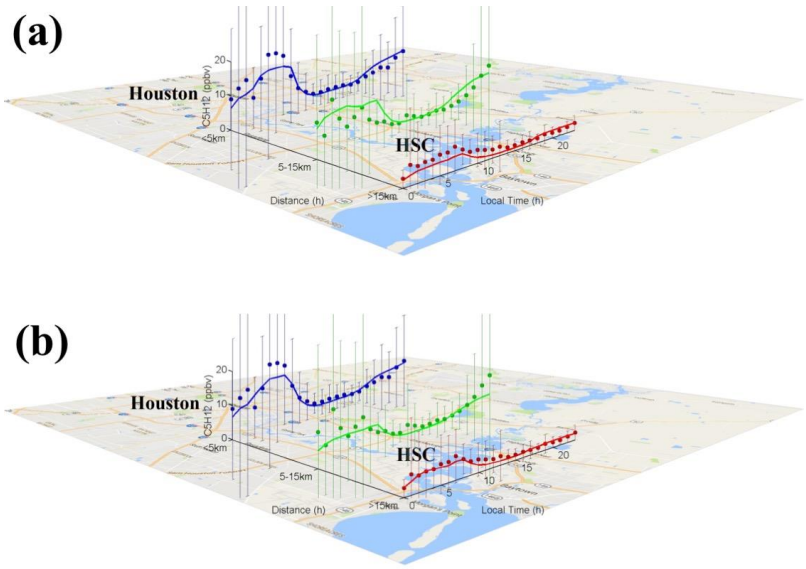

Figure 4.4 – Same as Figure 4.3 but for pentane.

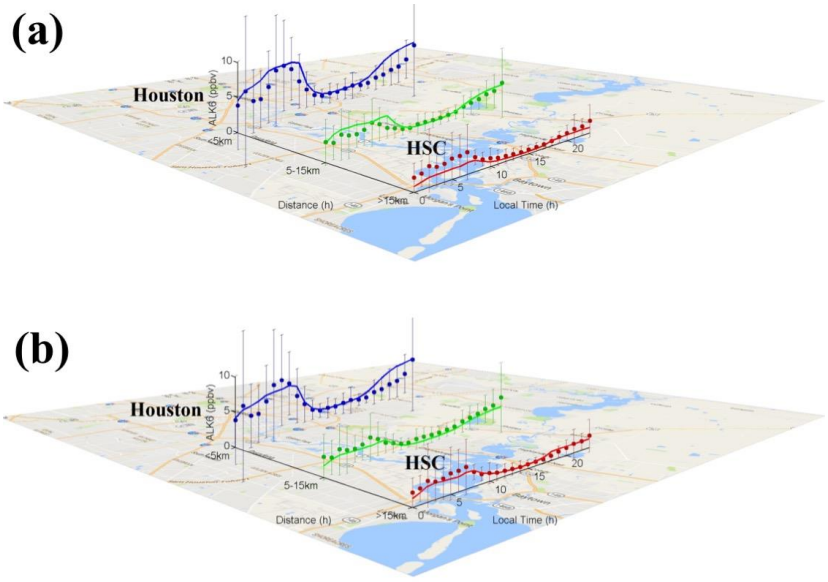

Figure 4.5 – Same as Figure 4.4 but for alkanes (C > 5).

4.2.3 VOCs Emissions Correction Scheme for Colorado

Alkanes simulated by REAM in Colorado are 50% to 70% lower than the FRAPPE measurements based on current emissions inventory of NEI 2011 (Figure 4.11). In Figure 4.12 we will show that O_3 is simulated ~20% lower than the observational data in the boundary layer. We suppose these two gaps are also the results of underestimations of VOCs emissions from O&NG operations like that of Houston. Similarly, we simplify the correction of emissions by choose a certain factor to be multiplied to the alkanes emissions on grid cells which cover the O&NG facilities in the research area according to current emissions inventory of NEI 2011. Emission sensitive analysis results indicates a factor of 5 leads to a better agreement of alkanes and O_3 between simulated and observed results. We will show the comparison in Figure 4.11 and Figure 4.12.

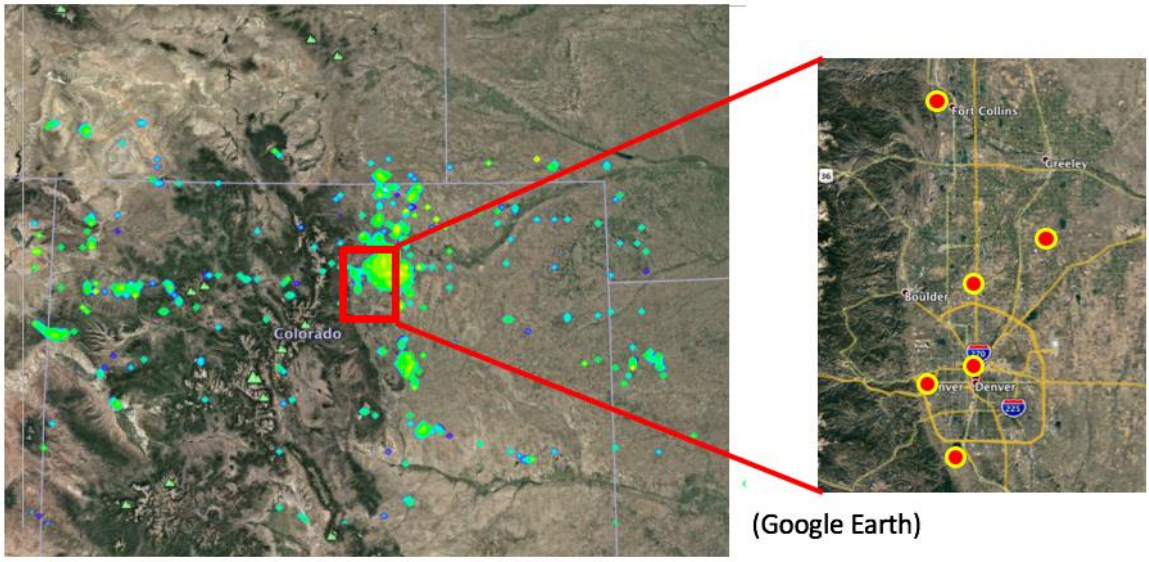

Figure 4.6 – Locations of O&NG facilities and the model domain of REAM (red dashed rectangles) which cover the former. Figure on left shows the distribution of flights spiral sites in DISCOVER-AQ 2014 campaign.

4.2.4 O_3-CH_2O regression slope analyse methods

To analyze the contribution from each source to the observed O_3-CH_2O relationship, we trace separately different CH_2O sources: primary anthropogenic emissions and the oxidation of anthropogenic VOCs ($CH_2O_{anthroVOCs}$), the oxidation of biogenic isoprene ($CH_2O_{bioISOP}$), and transport from model lateral and upper boundaries (CH_2O_{BC}). Other biogenic VOCs are not taken into account because isoprene provides the source for the vast majority of biogenic CH_2O [e.g., Kesselmeier and Staudt, 1999; Lathière et al., 2006; Guenther et al., 2012; Sindelarova et al., 2014]. In this tagged-tracer simulation, relevant species and radicals, such as O_3, NO_x, and HO_x (OH and HO_2), are specified with results archived from the standard simulation. The sum of the three individual tagged tracers is within 2% of the total CH_2O concentrations in the standard simulation. We carry out minor scaling adjustments, assuming that relative CH_2O attributions stay the same, in post-

processing to ensure that the sum of the CH_2O tracers is the same as the total CH_2O for each grid cell in the standard simulation. With simulated CH_2O attribution results, we can decompose the O_3-CH_2O regression slope into 3 sub-slopes of the corresponding CH_2O tracers (Eq. (9)). More detail explanations could be found in previous study [Cheng et al, 2017].

$$\text{Least} - \text{Squares Regression Slope of } O_3\text{-}CH_2O = \frac{Cov(CH_2O_{anthro}, O_3)}{Var(CH_2O_{total})} + \frac{Cov(CH_2O_{BC}, O_3)}{Var(CH_2O_{total})} + \frac{Cov(CH_2O_{bioISOP}, O_3)}{Var(CH_2O_{total})}, \quad (9)$$

In particular, the anthropogenic VOCs in Houston are further divided into 3 different types: the VOCs from the selected grid cells at HSC (CH_2O_{HSC}), VOCs from Houston urban areas (CH_2O_{urban}), and the rest of the VOCs from other anthropogenic sources ($CH_2O_{anthro\text{-}others}$) (Eq. (10)). And in Colorado, anthropogenic VOCs are divided into 2 different types: the VOCs from the oil and gas operations ($CH_2O_{oil\&gas}$), and the rest of the VOCs from other anthropogenic sources ($CH_2O_{anthro\text{-}others}$) (Eq. (11)).

$$\text{Least} - \text{Squares Regression Slope of } O_3\text{-}CH_2O_{Houston} = \frac{Cov(CH_2O_{HSC}, O_3)}{Var(CH_2O_{total})} + \frac{Cov(CH_2O_{urban}, O_3)}{Var(CH_2O_{total})} + \frac{Cov(CH_2O_{anthro-others}, O_3)}{Var(CH_2O_{total})} + \frac{Cov(CH_2O_{BC}, O_3)}{Var(CH_2O_{total})} + \frac{Cov(CH_2O_{bioISOP}, O_3)}{Var(CH_2O_{total})}, \quad (10)$$

$$\text{Least} - \text{Squares Regression Slope of } O_3\text{-}CH_2O_{\text{Colorado}}$$

$$= \frac{Cov(CH_2O_{oil\&gas}, O_3)}{Var(CH_2O_{\text{total}})} + \frac{Cov(CH_2O_{\text{anthro-others}}, O_3)}{Var(CH_2O_{\text{total}})} + \frac{Cov(CH_2O_{BC}, O_3)}{Var(CH_2O_{\text{total}})} + \frac{Cov(CH_2O_{\text{bioISOP}}, O_3)}{Var(CH_2O_{\text{total}})}, \quad (11)$$

where *Cov* and *Var* denote covariance and variance, respectively.

4.3 Results and Discussions

4.3.1 Simulations with Corrected Emissions on Houston

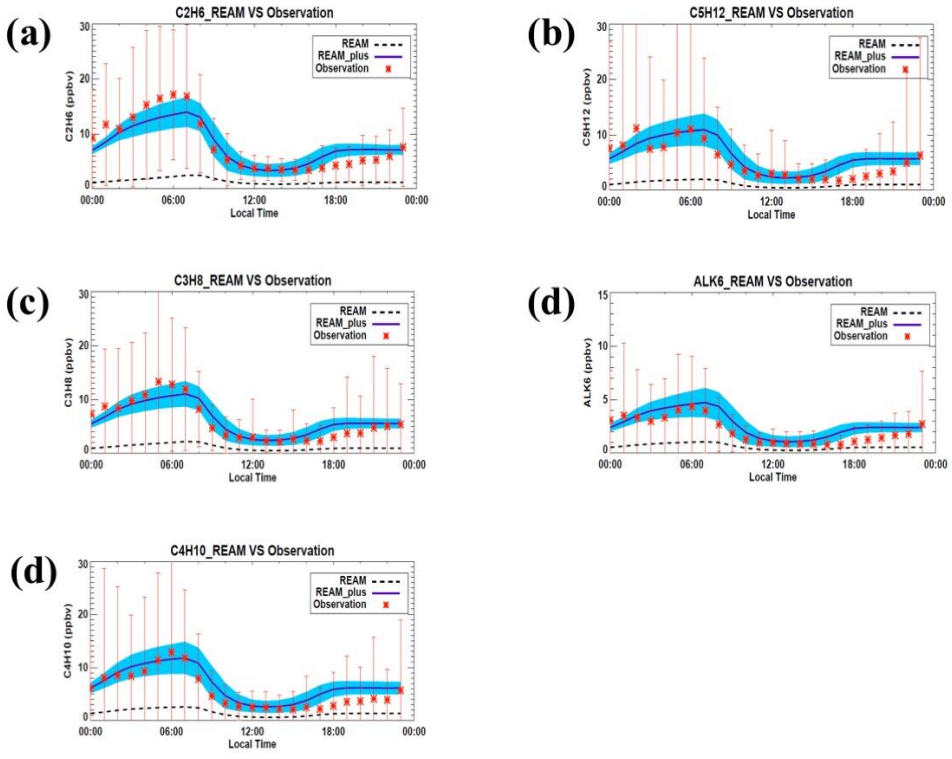

Figure 4.7 – Observed and simulated hourly mixing ratios of ethane (a), propane (b), butane (c), pentane (d) and alkanes (C > 5) (e) in September 2013. "REAM"

represents the original model results with of emission inventory of NEI 2011. "REAM_plus" represents the model results with increased alkanes emissions. The red horizontal bars show the observed standard deviations; the shaded blue areas denote the simulated standard deviations.

We first evaluate simulated hourly mixing ratios of alkanes with observational data. After increasing the alkanes emissions at the selected grid cells to 15 times of the original emissions of NEI 2011, the observational data are very well captured by the model simulations (Figure 4.7). Each one of the alkanes shows an increasing trend after the sunset (~ 6:00 pm) and reach a peak at ~6:00 am before sunrise in the absence of photochemistry sinks indicating a significant accumulation due to the high emissions. After sunrise, these alkanes decrease rapidly and reach their lowest mixing ratios at noon (~12:00 pm).

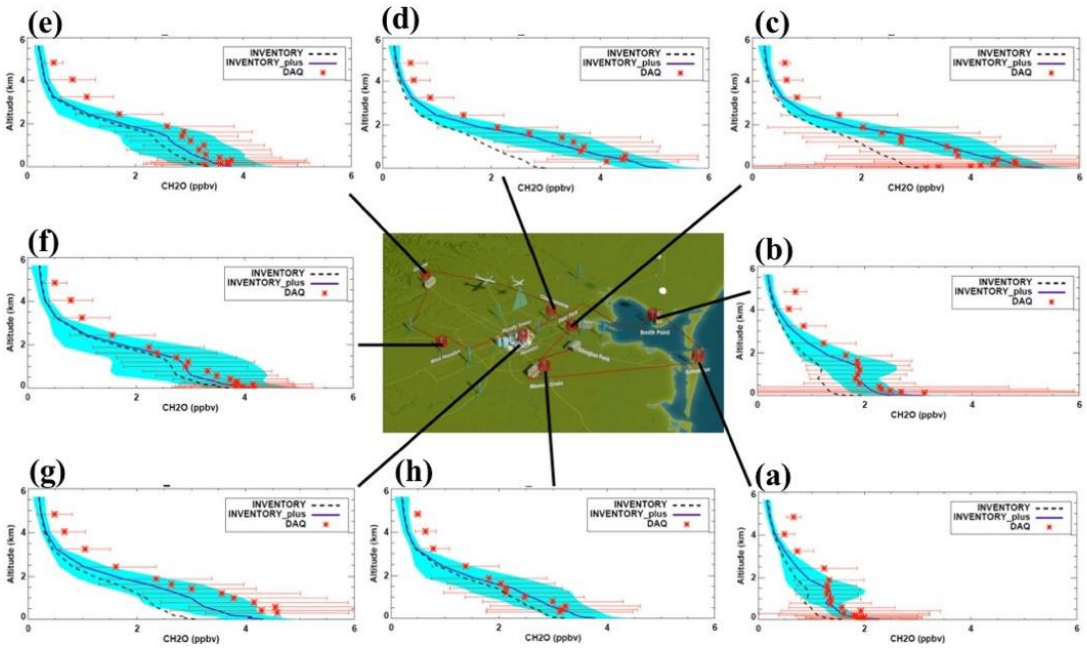

Figure 4.8 – Observed and simulated vertical profiles of CH_2O in 2013 during the DISCOVER-AQ experiment. DAQ represents DISCOVER-AQ data. The red horizontal bars show the observed standard deviations; the shaded blue areas denote the simulated standard deviations. "REAM" represents the original model results with of emission inventory of NEI 2011. "REAM_plus" represents the model results with increased alkanes emissions. The red horizontal bars show the observed

standard deviations; the shaded blue areas denote the simulated standard deviations. For comparison purposes, the observational data are binned vertically according to the vertical grid structure in the model.

We compare the simulated and observational vertical profiles of CH_2O in absence and attendance of the emissions correction (Figure 4.8). Multiplying 15 to the original alkanes emissions at the selected grids cells makes the simulated and observed results of CH_2O have good agreements in most spiral sites. In Houston urban site (g) and other two remote sites (e) and (f), there are still 5%-15% underestimation of CH_2O, indicating the existence of other primary or secondary sources at these areas. Except for the bay areas (a) and (b), the nearly linear decrease of CH_2O reflects that its atmospheric secondary (photochemical) source is much larger than primary emissions. For the sites (a) and (b), the vertical profiles of CH_2O are straightforward for 800 m to 2 km, reveal a well mixture in the boundary layer.

We also compered the simulated vertical profiles of O_3, CO, NO_x, and isoprene in Houston areas, with DISCOVER-AQ observations (Figure 4.9). The model reproduces well the observations. CO, with a longer life time, shows a nearly linear decreasing profile from 300 m to 5 km. Compared to the profile of CO in Maryland [Cheng, et al., 2017] which shows a large gradient near surface, the linear profile in Houston areas suggests that secondary formation of CO from oxidation of VOCs plays an important role. The short-lived NO_x and isoprene decrease rapidly from the surface to the free troposphere, reflecting in part the dominant NO_x and VOCs sources near the surface and rapid photochemical loss in the atmosphere. Ozone has a peak in the middle of the boundary layer due in part to dry deposition loss of ozone and in part to decreasing photochemical production and loss with altitude. It increases with altitude in the free troposphere reflecting the increase of chemical lifetime with altitude.

The daytime hourly variations of the selected species averaged from 300 m to 2.5 km are shown in Figure 4.10. In general, the variation is much less than in the vertical. The temporal change of NO_x is insignificant while O_3, CO, isoprene, and CH_2O tend to accumulate in the lower atmosphere from morning to afternoon. The observed variations are well simulated by the model, providing further evidence for the good model performance during the measurement period.

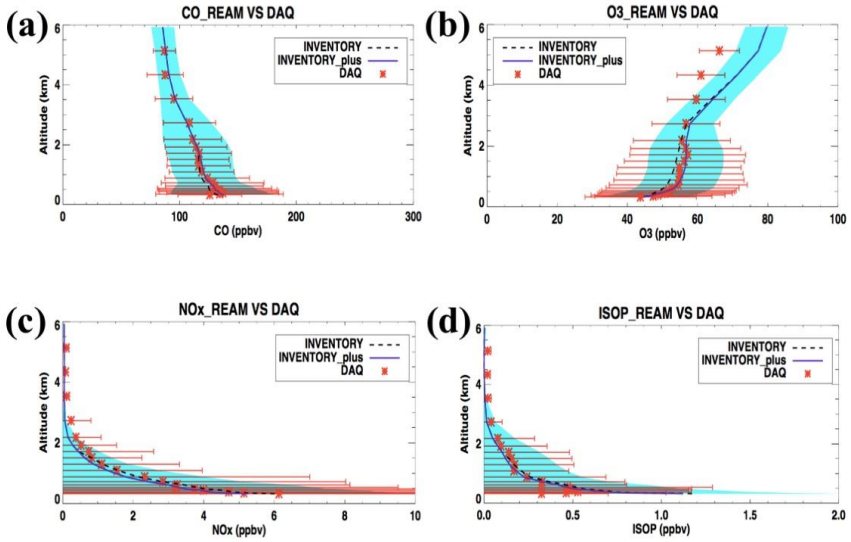

Figure 4.9 – Observed and simulated vertical profiles of CO (a), O_3 (b), NO (c), NO_2 (d) and isoprene (e) in 2013 during the DISCOVER-AQ experiment. Emissions are NEI 2011 with increased alkanes emissions. DAQ represents DISCOVER-AQ data. The red horizontal bars show the observed standard deviations; the shaded blue areas denote the simulated standard deviations. For comparison purposes, the observational data are binned vertically according to the vertical grid structure in the model.

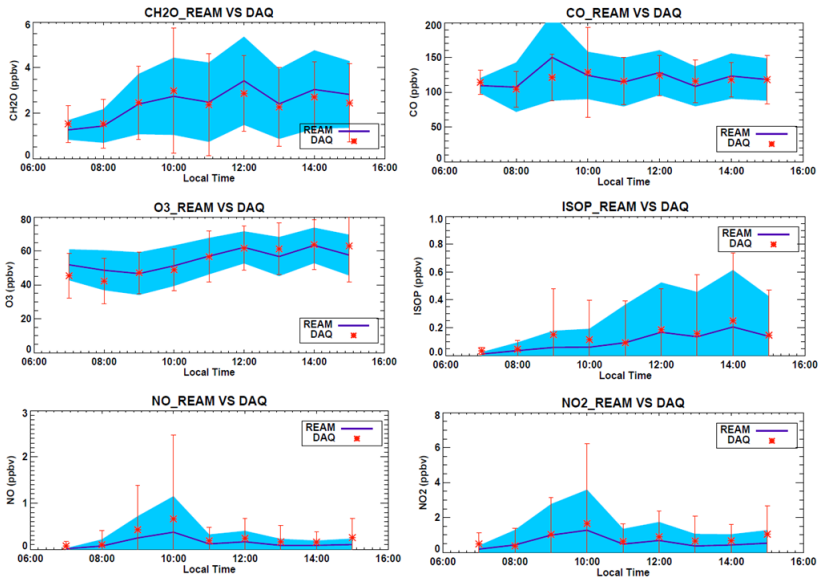

Figure 4.10 – Same as Figure 4.9 but for hourly trace gases mixing ratios. The red vertical bars show the observed standard deviations; the shaded blue areas denote the simulated standard deviations. Observed or corresponding simulated data from 300 m to 2.5 km are averaged hourly.

4.3.2 Simulations with Corrected Emissions on Colorado

We evaluate simulated profiles of the mixing ratios of alkanes in Colorado using FRAPPE observational data (Figure 4.11). Since there are no fixed spiral sites in FRAPPE measurements, model profiles are archived by grouping and averaging the data at model grid cells corresponding to the locations and time of aircraft samplings. Observational data are very well captured by the model by increasing the alkanes emissions at the O&NG identified grid cells to 5 times of the original emissions of NEI 2011.

We compare the simulated and observational vertical profiles of O_3 with and without emissions correction (Figure 4.12). In all 6 sites, observed O_3 shows obvious increase from surface to ~1 km, indicating the existence of abundant O_3 precursors. However, simulations with current NEI 2011 inventory do not show this trend. Multiplying

5 to the original alkanes emissions at the selected grids cells leads a significant improvement in simulation at most spiral sites. This result reveals that the simulated insufficient O_3 is due to the underestimation of alkanes emissions. Observed O_3 in site (1), (5) and (6) are well produced by model. In other sites, there are still 5%-10% underestimation of O_3. This indicates that even a constant factor of 5 gives a general amount of the underestimated alkanes emitted, for each individual site, more specific correction is still needed which requires more observations to provide the constraints.

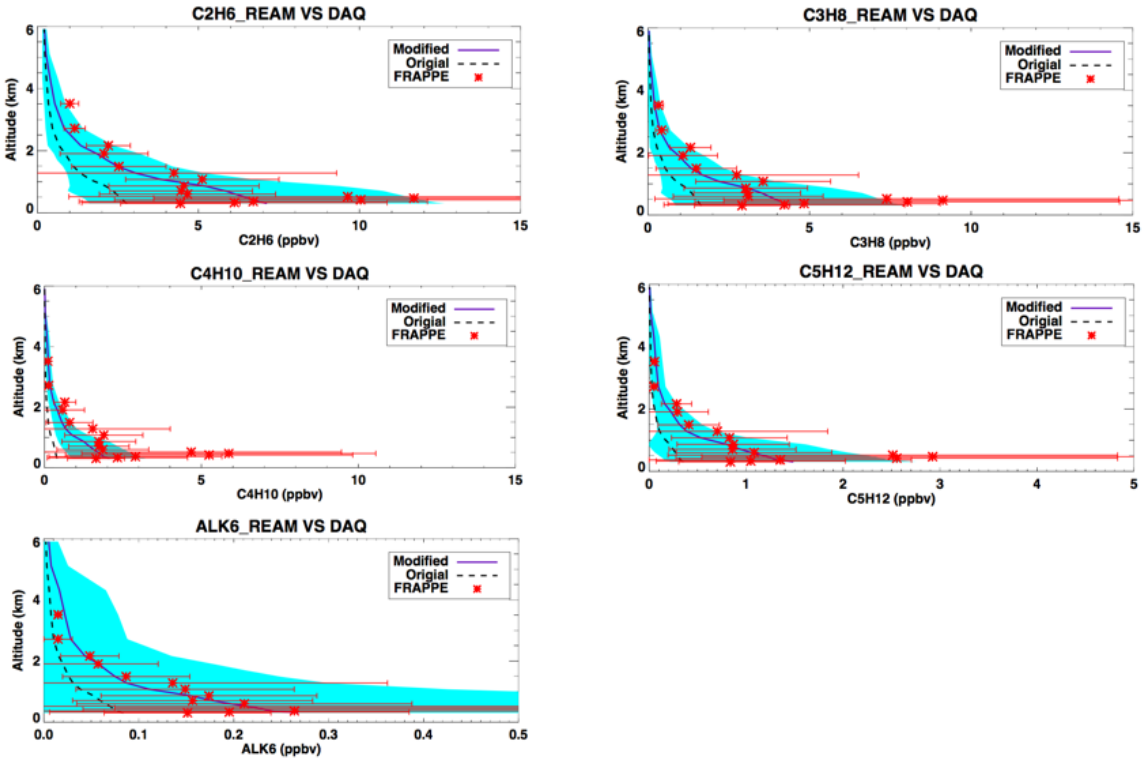

Figure 4.11 – Observed and simulated mixing ratios vertical profiles of ethane (a), propane (b), butane (c), pentane (d) and alkanes (C > 5) (e) in August 2014 Colorado sites. "Original" represents the original model results with of emission inventory of NEI 2011. "Modified" represents the model results with increased alkanes emissions. The red horizontal bars show the observed standard deviations; the shaded blue areas denote the simulated standard deviations.

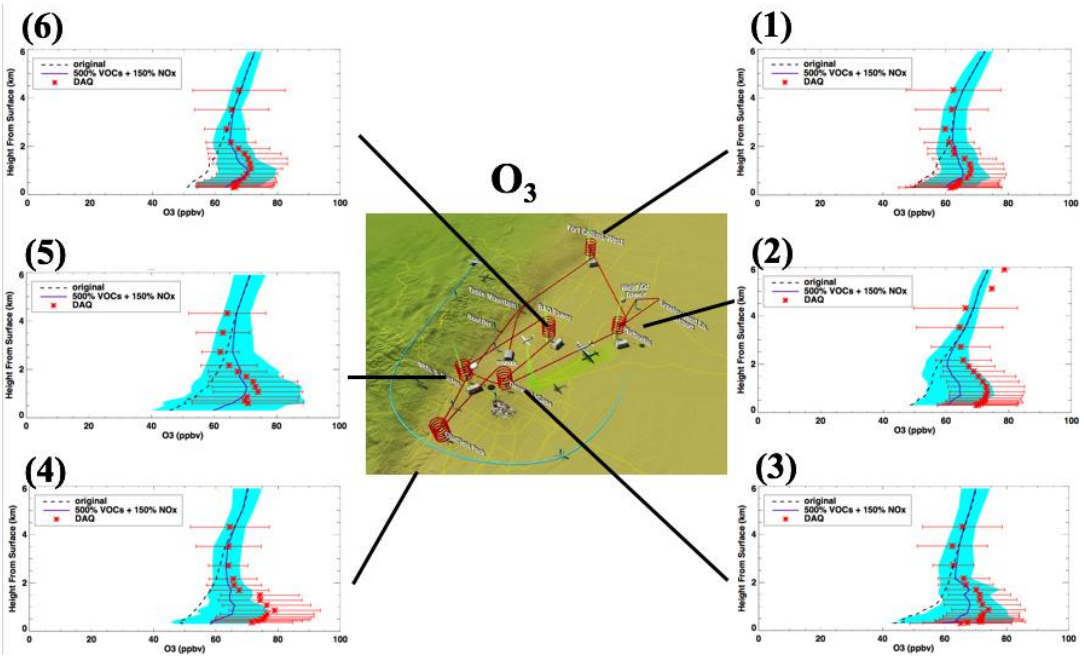

Figure 4.12 – Observed and simulated vertical profiles of O_3 in 2014 during the DISCOVER-AQ experiment. DAQ represents DISCOVER-AQ data. The red horizontal bars show the observed standard deviations; the shaded blue areas denote the simulated standard deviations. "REAM" represents the original model results with of emission inventory of NEI 2011. "REAM_plus" represents the model results with increased alkanes emissions. The red horizontal bars show the observed standard deviations; the shaded blue areas denote the simulated standard deviations. For comparison purposes, the observational data are binned vertically according to the vertical grid structure in the model.

We also compered the simulated vertical profiles of CH_2O, CO, and NO_x in 6 Colorado sites, with DISCOVER-AQ observations (Figure 4.13, Figure 4.14 and Figure 4.15). The model reproduces well the observations after emissions corrections.

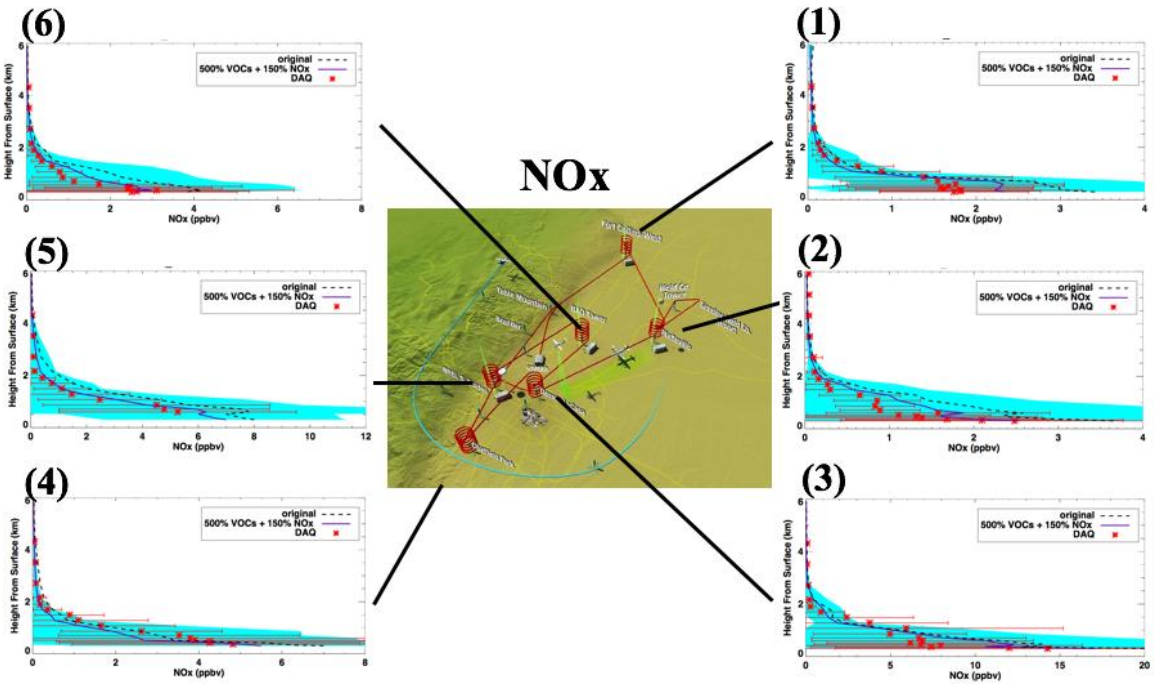

Figure 4.13 – Same as Figure 4.12 but for NO_x.

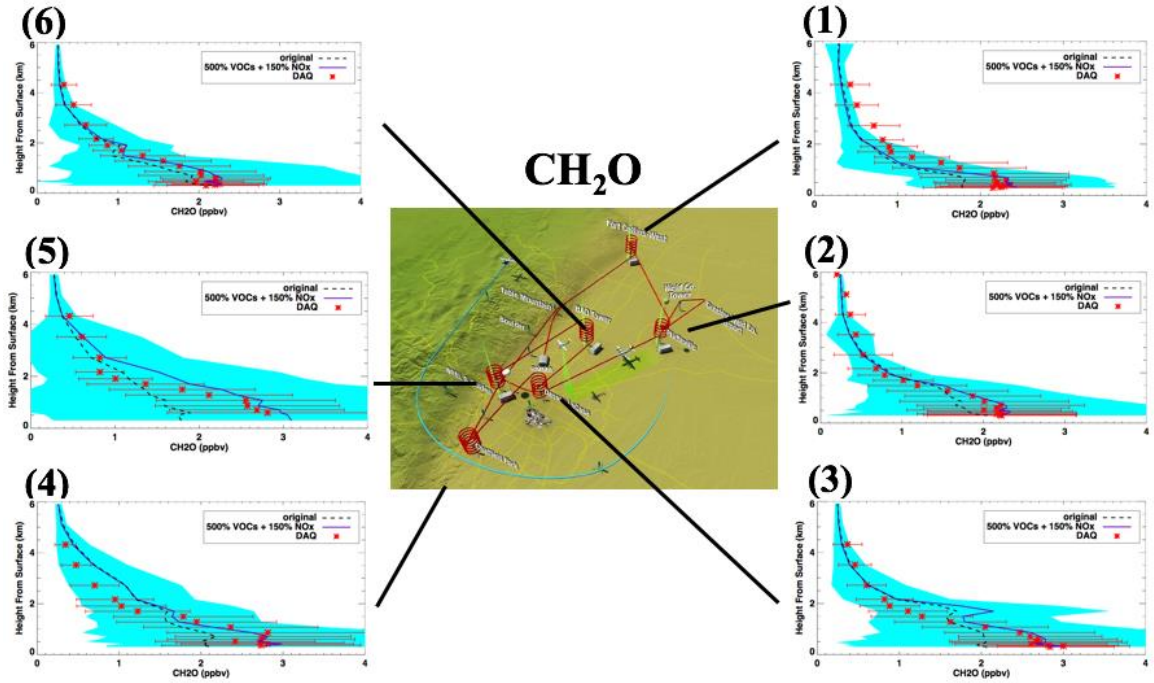

Figure 4.14 – Same as Figure 4.12 but for CH_2O.

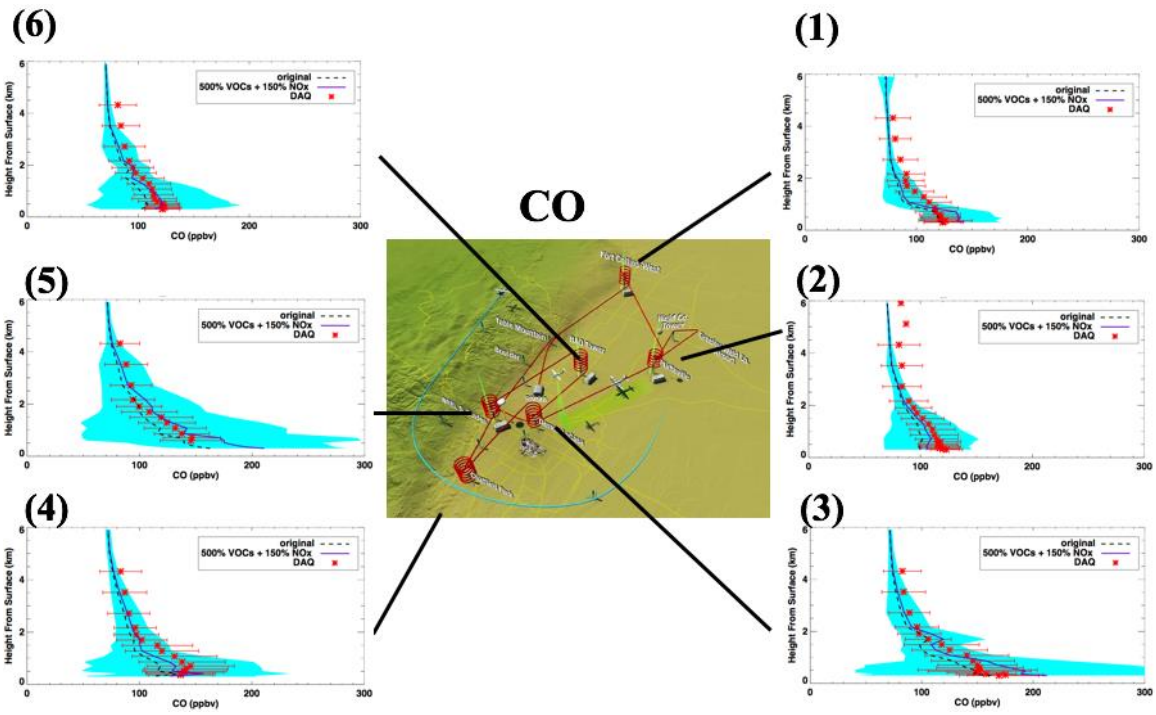

Figure 4.15 – Same as Figure 4.12 but for CO.

4.3.3 O_3-CH_2O regression slope

We compare simulated O_3-CH_2O correlations to the observations as a function of altitude or time of the day in Figure 4.16. The observations show a narrow range of R value (~0.5 to 0.75) from 300 m to the top of the boundary layer (~2 km). Below 300 m, the observations indicate relative lower R value (<0.5) down to surface. Similar to the R values of O_3-CH_2O in Maryland (Figure 3.1), in the free troposphere, CH_2O is dominated by the contribution by boundary layer transport and its magnitude level are very low, which is not related to O_3 production in the boundary layer. Near the surface, CH_2O is dominated by surface primary emissions, which are not directly related to O_3 production either. In the middle and upper boundary layer, the relative contribution by photochemical CH_2O

production increases. Therefore, the concurrent photochemical production of O_3 and CH_2O is likely the major factor contributing to the observed positive correlation between O_3 and CH_2O. The model results, even show a relative overestimate (~0.1) in R value from 300 m to the top of the boundary layer, are similar to the observed values and also show lower correlations near the surface and in the free troposphere. The slight bias between model and observational data is due to the correct is simplified by only increasing alkanes instead of all other VOCs. But in general, the strong agreement of model simulation and observation indicates that with the enhanced emissions of alkanes at selected grid cells, the O_3-CH_2O photochemistry is well captured. Figure 4.16 also shows the observed O_3-CH_2O correlation coefficient as a function of time of the day with data from 300 m to 2 km. This R value represents the spatial correlation in a given hour. It is somewhat lower than the temporal correlation discussed previously. The value ranges from 0 to 0.25 in the morning and ~0.5 in most of the afternoon. The model is in good agreement with the observations.

Figure 4.17 compares the observed and simulated O_3-CH_2O regression slopes in 3 different DISCOVER-AQ experiment campaigns. The O_3-CH_2O regression slope at a given altitude is simulated by the model with a satisfactory agreement in both Houston (a) and Denver (b). In Houston, above 2 km, the regression slope value decreases in the observation data but the model shows a decreasing regression slope value at 2.5 km. However, the observed decrease of the regression slope below 2 km is captured by the model. In Denver, both model and observation show bottom-up increasing slope below 1 km and decreasing slope above 1 km.

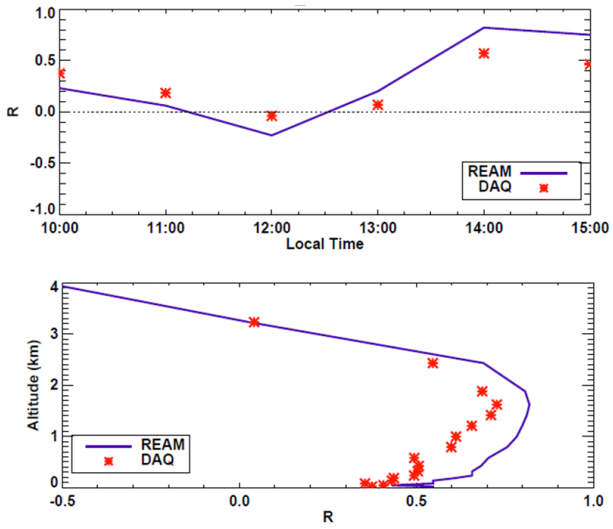

Figure 4.16 – Observed and simulated O_3-CH_2O correlation coefficients (R) as a function of altitude for daytime of 10 am - 4 pm (a) and local time for altitude of 0.3 - 2.5 km (b). The R values are computed using the observation or corresponding model data (for the entire DISCOVER-AQ experiment) at a given altitude bin or a given time period. Emissions are NEI 2011 with increased alkanes emissions.

We apply the model results to understand the relative contributions of CH_2O sources to the O_3-CH_2O regression slope. Figure 4.17 (a) shows that the contribution of CH_2O from biogenic isoprene oxidation in Houston is still the major one like that in Maryland (Figure 4.17 (c)). However, the contributions of CH_2O from photochemical oxidation of anthropogenic VOCs take place 40% of the total slope, which is different with that in Maryland. In particular, the emitted VOCs from selected grid cells at HSC contribute roughly 30% of total slope which reveals the ship-loading, transmit and storage of oil and gas is a significant contributor to the local atmospheric chemistry. The contribution of VOCs from Houston urban areas is about 10% of the total slope. Contributions from the model boundaries and other VOCs are minor in comparison. Compared with Maryland, the larger contribution from urban VOCs is consist with the suggestion from Zhu et al, [2014] that the VOCs is underestimated at this area. Different

from Houston and Maryland where isoprene oxidation plays the major role, Denver (Figure 4.17 (b)) shows that the O&NG contributes to ~80% of the total slope. Isoprene oxidation and other areas (including model boundaries) only take place ~10%, respectively. Though biogenic isoprene oxidation makes the largest contribution to the regression slope of O_3-CH_2O across much of the eastern United States (Figure 3.3), in the areas with heavy oil or gas operations like explore, transmit or storages, VOCs like alkanes from these facilities make comparative or even larger contributions to the regression slope of O_3-CH_2O. In these areas, since ozone variance is more correlated to alkanes, an urgent alkane emission control over the oil or gas operations is needed.

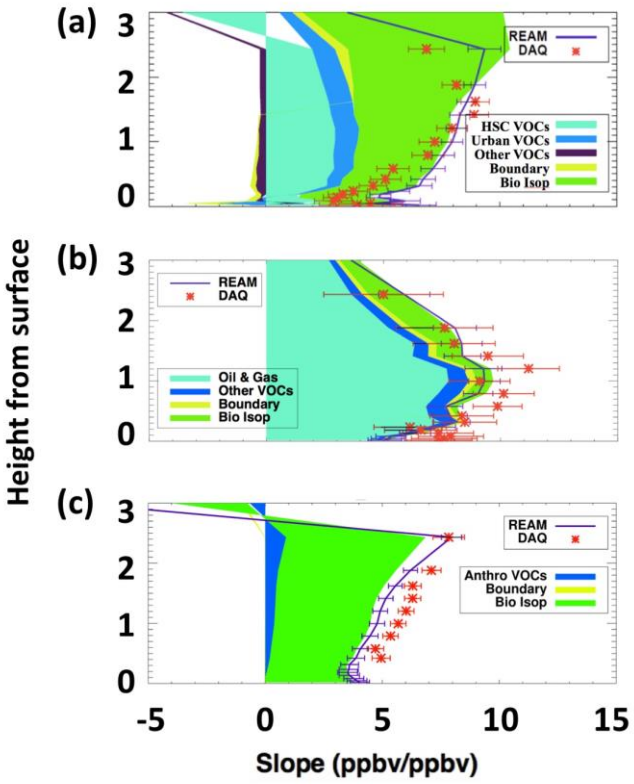

Figure 4.17 – Vertical profiles of O_3-CH_2O slope and sub slopes components in (a) September 2013 at Texas's DISCOVER AQ sites, (b) August 2014 at Colorado's DISCOVER AQ sites, and (c) July 2011 at Maryland's DISCOVER AQ sites. The

title "REAM" and "DAQ" stands for the results calculated by original total CH_2O and O_3 concentration in model and observation; In Houston, the titles of, "HSC VOCs", "Urban VOCs", "other VOCs", "Boundary" and "Bio Isop" stand for the results calculated by CH_2O from direct or indirect (oxidation of VOCs) from Houston ship channel emission, Houston urban emission, other kinds of anthropogenic emissions, boundary conditions, and biogenic isoprene's oxidation, respectively. In Colorado, the title of "Oil & Gas" stand for the results of emissions from oil and gas operations. In Maryland, the title "Anthro VOCs" represents the results from all anthropogenic VOCs emissions.

4.3.4 Summary and Implements

Table 4.1 – Comparison of VOCs emission correction factors proposed by different studies.

Areas	Species	Inventory	Increasing Factor	Methods	References
HSC	Alkanes	NEI2011	15	REAM	This study
HSC	VOCs	STARS[1]	5-15	SOF, DOAS	Johansson, 2014
Houston	HRVOCs	NEI05[2]	4.8±2.7	OMI	Zhu, 2014
Houston, HSC	Ethane, Propene	NEI05	7.66-27.58	WRF-Chem[4]	Kim, 2011
HSC	Alkanes	NEI05	8.7	SOF	Mellqvist, 2010
Houston	Ethene, Propene	TCEQ, SMOKE[3]	10	PEGASUS[4]	Jiang and Fast, 2004

[1] "STARS" donates the State of Texas Air Reporting System

[2] "NEI05" donates the State of Texas Air Reporting System

[3] "SMOKE" donates the State of Texas Air Reporting System

[4] "PEGASUS" donates the State of Texas Air Reporting System

The NASA DISCOVER-AQ airborne campaigns were carried out around the Houston and Denver metropolitan areas in the summers of 2013 and 2014, respectively. Using the 2011 national emissions inventory (NEI), REAM is applied to analyze the aircraft observations. We find that major model discrepancies are driven by large underestimates of alkane emissions in both regions. Modeling analysis suggests increases of alkane emissions by a factor of 15 in the Houston Ship Channel (Comparison of the results of this study and other studies are listed in Table 4.1), where ship-transport, ship-unloading, storage, domestic transportation of oil take place, and by a factor of 5 in the regions of oil and gas exploration of Denver. The large increase of alkane emissions has drastically different effects on O_3 concentrations depending on the strength of biogenic emissions. A useful metric to diagnose the effects of alkane emissions on photochemistry is the least-squares regression slope of O_3 to CH_2O, which increases by 30% and 80% in Houston and Denver, respectively, due to the increases of alkane emissions, leading to good agreement between model simulations and aircraft observations. Our finding implies that alkane emissions from oil and gas related sources may be substantially underestimated by the NEI, leading to corresponding underestimates of anthropogenic contributions to O_3 particularly over the western United States where biogenic VOC emissions are low. In regions like Denver, reducing alkane emissions is urgently required to control summertime O_3 concentrations.

CHAPTER 5. CONCLUSIONS AND FUTURE WORK

5.1 Summary of the findings

5.1.1 Biogenic and anthropogenic contributions to the regression slope of O_3-CO

The O_3-CO regression slope ~0.3 is simulated over the eastern outflow regions over the ocean. Over the eastern inland regions of both United States and China the O_3-CO regression slope is lower than that over the outflow region. At Maryland, O_3-CO regression slope in the boundary layer (300 m – 2.5 km) is controlled by both primary anthropogenic CO emissions and CO production by isoprene oxidation and the contribution by the latter is nearly as large as the former. For a larger scale, biogenic isoprene oxidation makes a significant contribution to the regression slope of O_3 versus CO in eastern United States both on the eastern continent and in pollutant outflow regions of oceans in summer, while in the western United States and China, the slope scenario is majorly anthropogenic CO emissions controlled.

Due to the continues decreasing trend of anthropogenic emissions, biogenic CO needs to be properly accounted for through modeling or measurement means for the eastern United States in the summer when using the O_3-CO regression slope and the total anthropogenic CO emissions to estimate the net anthropogenic O_3 production amount. The biogenic contribution will continue increasing as anthropogenic emissions decrease. Despite the change of the relative importance of anthropogenic and biogenic emissions, the consistent enhancement of O_3 relative to CO observed in the boundary layer, as indicated

by the O_3-CO regression slope, provides a useful constraint on model photochemistry and emissions.

5.1.2 Surface O_3 estimator with concentrations of CO and CH_2O

We build a surface ozone estimation model with linear regression approaches using given CH_2O and CO as input parameters, based on the found strong correlation and stable linear regression slope of O_3-CH_2O and CO-CH_2O. In this model， since biogenic isoprene oxidation makes most of the contribution to the regression slopes of CH_2O and O_3 in the southeast of United States, CH_2O is used to calculate O_3 and CO generated by the oxidation of biogenic VOCs. The estimator, with reasonable validation in the range of 70% to 150% of the anthropogenic emissions scenario in 2011, provides a fast and large scale surface O_3 estimation in southeast of United States in summer, given high-quality geostationary satellite observations of CO and CH_2O. The results of the estimation also indicate reducing anthropogenic NO_x emissions in southeast of United States can significantly decrease surface O_3 concentration.

5.1.3 VOCs emissions correction at oil and gas post-processed handing areas

Model simulations with US EPA inventory (2011) shows an obvious CH_2O underestimation by 20%-40%, and a O_3 underestimation by 20%-50%, indicating the existence of missing sources of VOCs in Houston and Denver, respectively. In this study, we apply a 3-D model with and VOCs correction by multiply a factor to the alkanes emissions of the model grid cells, selected based on the locations of post-processed handing, including ship unloading, storage, domestic or ship transportation of oil, at Houston ship channel and all the oil and gas operations in Denver. Our resulting estimate of alkanes emissions for Houston is 15 times higher and for Denver is 5times higher than the US EPA inventory. The simulation results with emissions corrections show a great agreement with

observation data of DISCOVER-AQ 2013 Taxes Champion in Houston and DISCOVER-AQ 2014 Colorado Champion in Denver, in CH_2O, NO_x, O_3, CO, and isoprene. The correlation analyze shows that oil and gas operations emission in Houston and Denver give about 30% and 80% contribution to the total correlation of CH_2O and O_3, respectively, which means oil and gas VOCs (alkanes) emission plays a significant role in local O_3 chemistry in these areas.

5.2 Recommendations for future work

5.2.1 Seasonal analyse on biogenic and anthropogenic contributions to slope of O_3-CO

In this dissertation, we have studied the contributions of biogenic and anthropogenic emissions to the slope of O_3-CO in summertime. However, the biogenic emissions have obvious seasonal variance which can change its proportion of the contribution. Besides, meteorology conditions, like solar radiations and temperature which are important factors in the O_3 production changes a lot in different seasons. Thus, the to conduct a longer term (covering different seasons) of model simulation will help us to better understand the characters of correlations of O_3-CO.

5.2.2 Performance of O_3 estimator in different conditions.

In this dissertation, the coefficients of O_3 estimator (Eq. (8)) is calculated using model results in July. As mentioned, correlations of O_3-CO, O_3-CH_2O and CO-CH_2O will be largely effected by the biogenic emissions and meteorology conditions. The performance of the estimator can be tested with the setup using model simulations at

different seasons. This work can help to study the implemental range of time of the estimator.

In addition, future work could focus on applying the estimator with high-quality geostationary satellite observations of CO and CH_2O to improve surface O_3 distribution estimation.

5.2.3 VOCs emissions correction scheme update.

The limitation of the emissions correction in this dissertation is that since the oil-gas operations are concentrated at HSC areas, there will be large uncertainty of attributing the missing sources to oil-gas post-processed handling, since the grid cells selected for corrections cannot entirely exclude other sources. The detail of the supplementary inventory of VOCs emissions at this area still need further measurements and model simulations study, for example, a higher resolution 3-D model could be used to identify the locations of selected grid cells.

APPENDIX A. DERIVATIONS OF EQUATIONS OF DECOMPOSITIONS OF O_3-CO AND O_3-CH_2O REGRESSION SLOPES

A.1 Derivations of Equation (1)

CO concentration is the sum of that from primary anthropogenic emissions ($CO_{anthroCO}$), the oxidation of anthropogenic VOCs ($CO_{anthroVOCs}$), the oxidation of biogenic isoprene ($CO_{bioISOP}$), and transport from model lateral and upper boundaries (CO_{BC}):

$$[\mathrm{CO}]_{\mathrm{total}} = [\mathrm{CO}]_{\mathrm{anthroCO}} + [\mathrm{CO}]_{\mathrm{anthroVOCs}} + [\mathrm{CO}]_{\mathrm{BC}} + [\mathrm{CO}]_{\mathrm{bioISOP}}. \quad \text{(A1)}$$

The slope of O_3 to CO in a least-squares regression is thus:

$$\mathrm{Least-Squares\ Regression\ Slope}$$

$$= \frac{Cov(\mathrm{CO}_{\mathrm{total}}, \mathrm{O}_3)}{Var(\mathrm{CO}_{\mathrm{total}})} \quad \text{(A2)}$$

$$= \frac{\overline{([\mathrm{CO}]_{\mathrm{total}} - \overline{[\mathrm{CO}]_{\mathrm{total}}})([\mathrm{O}_3] - \overline{[\mathrm{O}_3]})}}{Var(\mathrm{CO}_{\mathrm{total}})} \quad \text{(A3)}$$

$$= \frac{\overline{([\mathrm{CO}]_{\mathrm{anthroCO}} + [\mathrm{CO}]_{\mathrm{anthroVOCs}} + [\mathrm{CO}]_{\mathrm{BC}} + [\mathrm{CO}]_{\mathrm{bioISOP}} - \overline{([\mathrm{CO}]_{\mathrm{anthroCO}} + [\mathrm{CO}]_{\mathrm{anthroVOCs}} + [\mathrm{CO}]_{\mathrm{BC}} + [\mathrm{CO}]_{\mathrm{bioISOP}})})([\mathrm{O}_3] - \overline{[\mathrm{O}_3]})}}{Var(\mathrm{CO}_{\mathrm{total}})} \quad \text{(A4)}$$

$$= \frac{\overline{([\mathrm{CO}]_{\mathrm{anthroCO}} - \overline{[\mathrm{CO}]_{\mathrm{anthroCO}}})([\mathrm{O_3}] - \overline{[\mathrm{O_3}]})} + \overline{([\mathrm{CO}]_{\mathrm{anthroVOCs}} - \overline{[\mathrm{CO}]_{\mathrm{anthroVOCs}}})([\mathrm{O_3}] - \overline{[\mathrm{O_3}]})} + \overline{([\mathrm{CO}]_{\mathrm{BC}} - \overline{[\mathrm{CO}]_{\mathrm{BC}}})([\mathrm{O_3}] - \overline{[\mathrm{O_3}]})} + \overline{([\mathrm{CO}]_{\mathrm{bioISOP}} - \overline{[\mathrm{CO}]_{\mathrm{bioISOP}}})([\mathrm{O_3}] - \overline{[\mathrm{O_3}]})}}{Var(\mathrm{CO_{total}})}$$

(A5)

$$= \frac{Cov(\mathrm{CO_{anthroCO}}, \mathrm{O_3})}{Var(\mathrm{CO_{total}})} + \frac{Cov(\mathrm{CO_{anthroVOCs}}, \mathrm{O_3})}{Var(\mathrm{CO_{total}})} + \frac{Cov(\mathrm{CO_{BC}}, \mathrm{O_3})}{Var(\mathrm{CO_{total}})} + \frac{Cov(\mathrm{CO_{bioISOP}}, \mathrm{O_3})}{Var(\mathrm{CO_{total}})}, \quad \text{(A6)}$$

where $\bar{X}$ donates the average value of X.

A.2 Derivations of Equation (2)

CH_2O concentration is the sum of that from primary anthropogenic emissions and the oxidation of anthropogenic VOCs ($CH_2O_{anthroVOCs}$), the oxidation of biogenic isoprene ($CH_2O_{bioISOP}$), and transport from model lateral and upper boundaries (CH_2O_{BC}):

$$[\mathrm{CH_2O}]_{\mathrm{total}} = [\mathrm{CH_2O}]_{\mathrm{anthroVOCs}} + [\mathrm{CH_2O}]_{\mathrm{BC}} + [\mathrm{CH_2O}]_{\mathrm{bioISOP}}. \quad \text{(A1)}$$

The slope of O_3 to CH_2O in a least-squares regression is thus:

$$\text{Least} - \text{Squares Regression Slope of } \mathrm{O_3}\text{-}\mathrm{CH_2O} = \frac{Cov(\mathrm{CH_2O_{total}}, \mathrm{O_3})}{Var(\mathrm{CH_2O_{total}})} \quad \text{(A2)}$$

$$= \frac{\overline{([\mathrm{CH_2O}]_{\mathrm{total}} - \overline{[\mathrm{CH_2O}]_{\mathrm{total}}})([\mathrm{O_3}] - \overline{[\mathrm{O_3}]})}}{Var(\mathrm{CH_2O_{total}})} \quad \text{(A3)}$$

$$= \frac{\overline{([\mathrm{CH_2O}]_{\mathrm{anthroVOCs}} + [\mathrm{CH_2O}]_{\mathrm{BC}} + [\mathrm{CH_2O}]_{\mathrm{bioISOP}} - \overline{([\mathrm{CH_2O}]_{\mathrm{anthroVOCs}} + [\mathrm{CH_2O}]_{\mathrm{BC}} + [\mathrm{CH_2O}]_{\mathrm{bioISOP}})})([\mathrm{O_3}]}}{Var(\mathrm{CH_2O_{total}})}$$

(A4)

$$= \frac{\overline{\left([\mathrm{CH_2O}]_{\mathrm{anthroVOCs}} - \overline{[\mathrm{CH_2O}]_{\mathrm{anthroVOCs}}}\right)\left([\mathrm{O_3}] - \overline{[\mathrm{O_3}]}\right)}}{+\overline{\left([\mathrm{CH_2O}]_{\mathrm{BC}} - \overline{[\mathrm{CH_2O}]_{\mathrm{BC}}}\right)\left([\mathrm{O_3}] - \overline{[\mathrm{O_3}]}\right)} + \overline{\left([\mathrm{CH_2O}]_{\mathrm{bioISOP}} - \overline{[\mathrm{CH_2O}]_{\mathrm{bioISOP}}}\right)\left([\mathrm{O_3}] - \overline{[\mathrm{O_3}]}\right)}}{Var(\mathrm{CH_2O_{total}})} \quad \text{(A5)}$$

$$= \frac{Cov(\mathrm{CH_2O_{anthroVOCs}}, \mathrm{O_3})}{Var(\mathrm{CH_2O_{total}})} + \frac{Cov(\mathrm{CH_2O_{BC}}, \mathrm{O_3})}{Var(\mathrm{CH_2O_{total}})} + \frac{Cov(\mathrm{CH_2O_{bioISOP}}, \mathrm{O_3})}{Var(\mathrm{CH_2O_{total}})}, \quad \text{(A6)}$$

where $\bar{X}$ donates the average value of X.

REFERENCES

Atkinson, R., and J. Arey (1998), Atmospheric chemistry of biogenic organic compounds, Accounts of Chemical Research, 31(9), 574-583, doi:10.1021/ar970143z.

Atkinson, R., and J. Arey, Atmospheric chemistry of biogenic organic compounds, Accounts of Chemical Research, 31(9), 574-583, doi:10.1021/ar970143z, 1998.

Bak, J., J. Kim, X. Liu, K. Chance, and J. Kim, Evaluation of ozone profile and tropospheric ozone retrievals from GEMS and OMI spectra, Atmospheric Measurement Techniques, 6(2), 239, 2013.

Beer, R., TES on the Aura mission: Scientific objectives, measurements, and analysis overview, IEEE Transactions on Geoscience and remote sensing, 44(5), 1102-1105, 2006.

Bey, I., D. J. Jacob, R. M. Yantosca, J. A. Logan, B. D. Field, A. M. Fiore, Q. Li, H. Y. Liu, L. J. Mickley, and M. G. Schultz, Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation, Journal of Geophysical Research: Atmospheres, 106(D19), 23073-23095, doi:10.1029/2001JD000807, 2001.

Biancofiore, F., M. Verdecchia, P. Di Carlo, B. Tomassetti, E. Aruffo, M. Busilacchio, S. Bianco, S. Di Tommaso, and C. Colangeli, Analysis of surface ozone using a recurrent neural network, Science of the Total Environment, 514, 379-387, 2015.

Brent, L., W. Thorn, M. Gupta, B. Leen, J. Stehr, H. He, H. Arkinson, A. Weinheimer, C. Garland, and S. Pusede, Evaluation of the use of a commercially available cavity ringdown absorption spectrometer for measuring NO2 in flight, and observations over the Mid-Atlantic States, during DISCOVER-AQ, Journal of Atmospheric Chemistry, 72(3-4), 503-521, doi:10.1007/s10874-013-9265-6, 2015.

Buchwitz, M., I. Khlystova, H. Bovensmann, and J. Burrows, Three years of global carbon monoxide from SCIAMACHY: comparison with MOPITT and first results related to the detection of enhanced CO over cities, Atmospheric Chemistry and Physics, 7(9), 2399-2411, 2007.

Buhr, M., D. Sueper, M. Trainer, P. Goldan, B. Kuster, F. Fehsenfeld, G. Kok, R. Shillawski, and A. Schanot, Trace gas and aerosol measurements using aircraft data from the North Atlantic Regional Experiment (NARE 1993), Journal of Geophysical Research: Atmospheres, 101(D22), 29013-29027, doi:10.1029/96JD01159, 1996.

Burrows, W. R., M. Benjamin, S. Beauchamp, E. R. Lord, D. McCollor, and B. Thomson, CART decision-tree statistical analysis and prediction of summer season maximum surface ozone for the Vancouver, Montreal, and Atlantic regions of Canada, Journal of applied meteorology, 34(8), 1848-1862, 1995.

Buzcu, B., and M. P. Fraser, Source identification and apportionment of volatile organic compounds in Houston, TX, Atmospheric Environment, 40(13), 2385-2400, 2006.

Cardenas, L., J. Austin, R. Burgess, K. Clemitshaw, S. Dorling, S. Penkett, and R. Harrison, Correlations between CO, NO y, O 3 and non-methane hydrocarbons and their relationships with meteorology during winter 1993 on the North Norfolk Coast, UK, Atmospheric Environment, 32(19), 3339-3351, doi:10.1016/S1352-2310(97)00445-7, 1998.

Chameides, W., R. Lindsay, J. Richardson, and C. Kiang, The role of biogenic hydrocarbons in urban photochemical smog: Atlanta as a case study, Science, 241(4872), 1473-1475, 1988.

Chance, K., X. Liu, R. M. Suleiman, D. E. Flittner, J. Al-Saadi, and S. J. Janz (2013), Tropospheric emissions: monitoring of pollution (TEMPO), paper presented at Earth Observing Systems XVIII, International Society for Optics and Photonics.

Chance, K., P. I. Palmer, R. J. Spurr, R. V. Martin, T. P. Kurosu, and D. J. Jacob, Satellite observations of formaldehyde over North America from GOME, Geophysical Research Letters, 27(21), 3461-3464, 2000.

Cheng, Y., Y. Wang, Y. Zhang, G. Chen, J. H. Crawford, M. M. Kleb, G. S. Diskin, and A. J. Weinheimer, Large biogenic contribution to boundary layer O3-CO regression slope in summer, Geophysical Research Letters, 2017.

Chin, M., D. J. Jacob, J. W. Munger, D. D. Parrish, and B. G. Doddridge, Relationship of ozone and carbon monoxide over North America, Journal of Geophysical Research: Atmospheres, 99(D7), 14565-14573, doi:10.1029/94JD00907, 1994.

Choi, Y., G. Osterman, A. Eldering, Y. Wang, and E. Edgerton, Understanding the contributions of anthropogenic and biogenic sources to CO enhancements and outflow observed over North America and the western Atlantic Ocean by TES and MOPITT, Atmospheric Environment, 44(16), 2033-2042, doi:10.1016/j.atmosenv.2010.01.029, 2010.

Choi, Y., Y. Wang, Q. Yang, D. Cunnold, T. Zeng, C. Shim, M. Luo, A. Eldering, E. Bucsela, and J. Gleason, Spring to summer northward migration of high O3 over the western North Atlantic, Geophysical Research Letters, 35(4), doi:10.1029/2007GL032276, 2008.

Choi, Y., Y. Wang, T. Zeng, D. Cunnold, E. S. Yang, R. Martin, K. Chance, V. Thouret, and E. Edgerton, Springtime transitions of NO2, CO, and O3 over North America: Model evaluation and analysis, Journal of Geophysical Research: Atmospheres, 113(D20), doi:10.1029/2007JD009632, 2008.

Choi, Y., Y. Wang, T. Zeng, R. V. Martin, T. P. Kurosu, and K. Chance, Evidence of lightning NOx and convective transport of pollutants in satellite observations over North America, Geophysical Research Letters, 32(2), doi:10.1029/2004GL021436, 2005.

Cobourn, W. G., Accuracy and reliability of an automated air quality forecast system for ozone in seven Kentucky metropolitan areas, Atmospheric Environment, 41(28), 5863-5875, 2007.

Cooper, O., J. Moody, D. Parrish, M. Trainer, J. Holloway, G. Hübler, F. Fehsenfeld, and A. Stohl, Trace gas composition of midlatitude cyclones over the western North Atlantic Ocean: A seasonal comparison of O3 and CO, Journal of Geophysical Research: Atmospheres, 107(D7), doi:10.1029/2001JD000902, 2002.

Cooper, O., J. Moody, D. Parrish, M. Trainer, T. Ryerson, J. Holloway, G. Hübler, F. Fehsenfeld, and M. Evans, Trace gas composition of midlatitude cyclones over the western North Atlantic Ocean: A conceptual model, Journal of Geophysical Research: Atmospheres, 107(D7), doi:10.1029/2001JD000901, 2002.

Cuesta, J., M. Eremenko, X. Liu, G. Dufour, Z. Cai, M. Höpfner, T. v. Clarmann, P. Sellitto, G. Forêt, and B. Gaubert, Satellite observation of lowermost tropospheric ozone by multispectral synergism of IASI thermal infrared and GOME-2 ultraviolet measurements over Europe, Atmospheric Chemistry and Physics, 13(19), 9675-9693, 2013.

Dabberdt, W. F., M. A. Carroll, D. Baumgardner, G. Carmichael, R. Cohen, T. Dye, J. Ellis, G. Grell, S. Grimmond, and S. Hanna, Meteorological research needs for improved air quality forecasting: Report of the 11th Prospectus Development Team of the US Weather Research Program, Bulletin of the American Meteorological Society, 85(4), 563-586, 2004.

Dasgupta, P. K., J. Li, G. Zhang, W. T. Luke, W. A. McClenny, J. Stutz, and A. Fried, Summertime ambient formaldehyde in five US metropolitan areas: Nashville, Atlanta, Houston, Philadelphia, and Tampa, Environmental science & technology, 39(13), 4767-4783, 2005.

Daum, P., L. Kleinman, L. Newman, W. Luke, J. Weinstein-Lloyd, C. Berkowitz, and K. Busness, Chemical and physical properties of plumes of anthropogenic pollutants transported over the North Atlantic during the North Atlantic Regional Experiment, Journal of Geophysical Research: Atmospheres, 101(D22), 29029-29042, 1996.

De Gouw, J., S. te Lintel Hekkert, J. Mellqvist, C. Warneke, E. Atlas, F. Fehsenfeld, A. Fried, G. Frost, F. Harren, and J. Holloway, Airborne measurements of ethene from industrial sources using laser photo-acoustic spectroscopy, Environmental science & technology, 43(7), 2437-2442, 2009.

De Laat, A., R. Dijkstra, H. Schrijver, P. Nédélec, and I. Aben, Validation of six years of SCIAMACHY carbon monoxide observations using MOZAIC CO profile measurements, Atmospheric Measurement Techniques, 5(9), 2133-2142, 2012.

De Smedt, I., M. Van Roozendael, T. Stavrakou, J. Müller, C. Lerot, N. Theys, P. Valks, N. Hao, and R. Van Der A, Improved retrieval of global tropospheric formaldehyde columns from GOME-2/MetOp-A addressing noise reduction and instrumental degradation issues, Atmospheric Measurement Techniques, 5(11), 2933-2949, 2012.

Delle Monache, L., X. Deng, Y. Zhou, and R. Stull, Ozone ensemble forecasts: 1. A new ensemble design, Journal of Geophysical Research: Atmospheres, 111(D5), 2006.

Delle Monache, L., J. P. Hacker, Y. Zhou, X. Deng, and R. B. Stull, Probabilistic aspects of meteorological and ozone regional ensemble forecasts, Journal of Geophysical Research: Atmospheres, 111(D24), 2006.

Duane, M., B. Poma, D. Rembges, C. Astorga, and B. Larsen, Isoprene and its degradation products as strong ozone precursors in Insubria, Northern Italy, Atmospheric Environment, 36(24), 3867-3879, 2002.

Duncan, B., J. Logan, I. Bey, I. Megretskaia, R. Yantosca, P. Novelli, N. B. Jones, and C. Rinsland, Global budget of CO, 1988–1997: Source estimates and validation with a global model, Journal of Geophysical Research: Atmospheres, 112(D22), 2007.

Emmons, L., M. Deeter, J. Gille, D. Edwards, J. L. Attié, J. Warner, D. Ziskin, G. Francis, B. Khattatov, and V. Yudin, Validation of Measurements of Pollution in the Troposphere (MOPITT) CO retrievals with aircraft in situ profiles, Journal of Geophysical Research: Atmospheres, 109(D3), 2004.

Friedfeld, S., M. Fraser, K. Ensor, S. Tribble, D. Rehle, D. Leleux, and F. Tittel, Statistical analysis of primary and secondary atmospheric formaldehyde, Atmospheric Environment, 36(30), 4767-4775, 2002.

Geng, F., X. Tie, A. Guenther, G. Li, J. Cao, and P. Harley, Effect of isoprene emissions from major forests on ozone formation in the city of Shanghai, China, Atmospheric Chemistry and Physics, 11(20), 10449-10459, doi:10.5194/acp-11-10449-2011, 2011.

Gilman, J. B., B. M. Lerner, W. C. Kuster, and J. De Gouw, Source signature of volatile organic compounds from oil and natural gas operations in northeastern Colorado, Environmental science & technology, 47(3), 1297-1305, 2013.

Gloudemans, A., H. Schrijver, Q. Kleipool, M. Van den Broek, A. Straume, G. Lichtenberg, R. Van Hees, I. Aben, and J. Meirink, The impact of SCIAMACHY near-infrared instrument calibration on CH 4 and CO total columns, Atmospheric Chemistry and Physics, 5(9), 2369-2383, 2005.

González Abad, G., X. Liu, K. Chance, H. Wang, T. Kurosu, and R. Suleiman, Updated Smithsonian Astrophysical Observatory Ozone Monitoring Instrument (SAO OMI) formaldehyde retrieval, Atmospheric Measurement Techniques, 8(1), 19-32, 2015.

González Abad, G., A. Vasilkov, C. Seftor, X. Liu, and K. Chance, Smithsonian Astrophysical Observatory Ozone Mapping and Profiler Suite (SAO OMPS) formaldehyde retrieval, Atmospheric Measurement Techniques, 9(7), 2797-2812, 2016.

Gu, D., Y. Wang, C. Smeltzer, and K. F. Boersma, Anthropogenic emissions of NOx over China: Reconciling the difference of inverse modeling results using GOME-2 and OMI measurements, Journal of Geophysical Research: Atmospheres, 119(12), 7732-7740, doi:10.1002/2014JD021644, 2014.

Gu, D., Y. Wang, C. Smeltzer, and Z. Liu, Reduction in NO x emission trends over China: Regional and seasonal variations, Environmental science & technology, 47(22), 12912-12919, doi:10.1021/es401727e, 2013.

Gu, D., Y. Wang, R. Yin, Y. Zhang, and C. Smeltzer, Inverse modelling of NO x emissions over eastern China: uncertainties due to chemical non-linearity, Atmospheric Measurement Techniques, 9(10), 5193, doi:10.5194/amt-9-5193-2016, 2016.

Guenther, A., C. N. Hewitt, D. Erickson, R. Fall, C. Geron, T. Graedel, P. Harley, L. Klinger, M. Lerdau, and W. McKay, A global model of natural volatile organic compound emissions, Journal of Geophysical Research: Atmospheres, 100(D5), 8873-8892, doi:10.1029/94JD02950, 1995.

Guenther, A., X. Jiang, C. Heald, T. Sakulyanontvittaya, T. Duhl, L. Emmons, and X. Wang, The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2. 1): an extended and updated framework for modeling biogenic emissions, 2012.

Guillas, S., J. Bao, Y. Choi, and Y. Wang, Statistical correction and downscaling of chemical transport model ozone forecasts over Atlanta, Atmospheric Environment, 42(6), 1338-1348, 2008.

Guven, B. B., and E. P. Olaguer, Ambient formaldehyde source attribution in Houston during TexAQS II and TRAMP, Atmospheric environment, 45(25), 4272-4280, 2011.

Hassler, B., B. C. McDonald, G. J. Frost, A. Borbon, D. C. Carslaw, K. Civerolo, C. Granier, P. S. Monks, S. Monks, and D. D. Parrish, Analysis of long-term observations of NOx and CO in megacities and application to constraining emissions inventories, Geophysical Research Letters, 43(18), 9920-9930, 2016.

Holloway, T., H. Levy, and P. Kasibhatla, Global distribution of carbon monoxide, Journal of Geophysical Research: Atmospheres, 105(D10), 12123-12147, 2000.

Honrath, R., R. C. Owen, M. Val Martin, J. Reid, K. Lapina, P. Fialho, M. P. Dziobak, J. Kleissl, and D. Westphal, Regional and hemispheric impacts of anthropogenic and biomass burning emissions on summertime CO and O3 in the North Atlantic lower free troposphere, Journal of Geophysical Research: Atmospheres, 109(D24), doi:10.1029/2004JD005147, 2004.

Hudman, R. C., L. T. Murray, D. J. Jacob, D. Millet, S. Turquety, S. Wu, D. Blake, A. Goldstein, J. Holloway, and G. Sachse, Biogenic versus anthropogenic sources of CO in the United States, Geophysical Research Letters, 35(4), doi:10.1029/2007GL032393, 2008.

Huntrieser, H., J. Heland, H. Schlager, C. Forster, A. Stohl, H. Aufmhoff, F. Arnold, H. Scheel, M. Campana, and S. Gilge, Intercontinental air pollution transport from North America to Europe: Experimental evidence from airborne measurements and surface observations, Journal of Geophysical Research: Atmospheres, 110(D1), doi:10.1029/2004JD005045, 2005.

Ingmann, P., B. Veihelmann, J. Langen, D. Lamarre, H. Stark, and G. B. Courrèges-Lacoste, Requirements for the GMES Atmosphere Service and ESA's implementation concept: Sentinels-4/-5 and-5p, Remote Sensing of Environment, 120, 58-69, 2012.

Jing, P., D. Cunnold, Y. Choi, and Y. Wang, Summertime tropospheric ozone columns from Aura OMI/MLS measurements versus regional model results over the United States, Geophysical research letters, 33(17), doi:10.1029/2006GL026473, 2006.

Jobson, B., C. M. Berkowitz, W. Kuster, P. Goldan, E. Williams, F. Fesenfeld, E. Apel, T. Karl, W. A. Lonneman, and D. Riemer, Hydrocarbon source signatures in Houston, Texas: Influence of the petrochemical industry, Journal of Geophysical Research: Atmospheres, 109(D24), 2004.

Johansson, J. K., J. Mellqvist, J. Samuelsson, B. Offerle, B. Lefer, B. Rappenglück, J. Flynn, and G. Yarwood, Emission measurements of alkenes, alkanes, SO2, and NO2 from stationary sources in Southeast Texas over a 5 year period using SOF and mobile DOAS, Journal of Geophysical Research: Atmospheres, 119(4), 1973-1991, 2014.

Johansson, J. K., J. Mellqvist, J. Samuelsson, B. Offerle, J. Moldanova, B. Rappenglück, B. Lefer, and J. Flynn, Quantitative measurements and modeling of industrial formaldehyde emissions in the Greater Houston area during campaigns in 2009 and 2011, Journal of Geophysical Research: Atmospheres, 119(7), 4303-4322, 2014.

Karl, T., T. Jobson, W. C. Kuster, E. Williams, J. Stutz, R. Shetter, S. R. Hall, P. Goldan, F. Fehsenfeld, and W. Lindinger, Use of proton-transfer-reaction mass spectrometry to characterize volatile organic compound sources at the La Porte super site during the Texas Air Quality Study 2000, Journal of Geophysical Research: Atmospheres, 108(D16), 2003.

Kesselmeier, J., L. Schäfer, P. Ciccioli, E. Brancaleoni, A. Cecinato, M. Frattoni, P. Foster, V. Jacob, J. Denis, and J. Fugit, Emission of monoterpenes and isoprene from a Mediterranean oak species Quercus ilex L. measured within the BEMA (Biogenic Emissions in the Mediterranean Area) project, Atmospheric Environment, 30(10-11), 1841-1850, 1996.

Kesselmeier, J., and M. Staudt, Biogenic volatile organic compounds (VOC): an overview on emission, physiology and ecology, Journal of atmospheric chemistry, 33(1), 23-88, 1999.

Khalil, M., and R. Rasmussen, Global decrease in atmospheric carbon monoxide concentration, Nature, 370(6491), 639, 1994.

Kim, E., P. K. Hopke, and E. S. Edgerton, Source identification of Atlanta aerosol by positive matrix factorization, Journal of the Air & Waste Management Association, 53(6), 731-739, 2003.

Kleinman, L. I., P. Daum, D. Imre, Y. N. Lee, L. Nunnermacker, S. Springston, J. Weinstein-Lloyd, and J. Rudolph, Ozone production rate and hydrocarbon reactivity in 5 urban areas: A cause of high ozone concentration in Houston, Geophysical Research Letters, 29(10), 2002.

Koo, J.-H., Y. Wang, T. Kurosu, K. Chance, A. Rozanov, A. Richter, S. Oltmans, A. Thompson, J. Hair, and M. Fenn, Characteristics of tropospheric ozone depletion events in the Arctic spring: analysis of the ARCTAS, ARCPAC, and ARCIONS measurements and satellite BrO observations, Atmospheric Chemistry and Physics, 12(20), 9909-9922, doi:10.5194/acp-12-9909-2012, 2012.

Kurosu, T. P., K. Chance, and C. E. Sioris (2004), Preliminary results for HCHO and BrO from the EOS-aura ozone monitoring instrument, paper presented at Passive Optical Remote Sensing of the Atmosphere and Clouds IV, International Society for Optics and Photonics.

Lathiere, J., D. Hauglustaine, A. Friend, N. D. Noblet-Ducoudré, N. Viovy, and G. Folberth, Impact of climate variability and land use changes on global biogenic volatile organic compound emissions, Atmospheric Chemistry and Physics, 6(8), 2129-2146, 2006.

Lee, K.-Y., K.-H. Kwak, Y.-H. Ryu, S.-H. Lee, and J.-J. Baik, Impacts of biogenic isoprene emission on ozone air quality in the Seoul metropolitan area, Atmospheric Environment, 96, 209-219, doi:10.1016/j.atmosenv.2014.07.036, 2014.

Lelieveld, J., and F. J. Dentener, What controls tropospheric ozone?, Journal of Geophysical Research: Atmospheres, 105(D3), 3531-3551, doi:10.1029/1999JD901011, 2000.

Leuchner, M., and B. Rappenglück, VOC source–receptor relationships in Houston during TexAQS-II, Atmospheric Environment, 44(33), 4056-4067, 2010.

Li, C., J. Joiner, N. A. Krotkov, and L. Dunlap, A new method for global retrievals of HCHO total columns from the Suomi National Polar-orbiting Partnership Ozone Mapping and Profiler Suite, Geophysical Research Letters, 42(7), 2515-2522, 2015.

Li, Q., D. J. Jacob, I. Bey, P. I. Palmer, B. N. Duncan, B. D. Field, R. V. Martin, A. M. Fiore, R. M. Yantosca, and D. D. Parrish, Transatlantic transport of pollution and its effects on surface ozone in Europe and North America, Journal of Geophysical Research: Atmospheres, 107(D13), doi:10.1029/2001JD001422, 2002.

Lindinger, W., and A. Jordan, Proton-transfer-reaction mass spectrometry (PTR–MS): on-line monitoring of volatile organic compounds at pptv levels, Chemical Society Reviews, 27(5), 347-375, 1998.

Liu, Z., Y. Wang, F. Costabile, A. Amoroso, C. Zhao, L. G. Huey, R. Stickel, J. Liao, and T. Zhu, Evidence of aerosols as a media for rapid daytime HONO production over China, Environmental science & technology, 48(24), 14386-14391, doi:10.1021/es504163z, 2014.

Liu, Z., Y. Wang, D. Gu, C. Zhao, L. Huey, R. Stickel, J. Liao, M. Shao, T. Zhu, and L. Zeng, Summertime photochemistry during CAREBeijing-2007: RO x budgets and O 3 formation, Atmospheric Chemistry and Physics, 12(16), 7737-7752, doi:10.5194/acp-12-7737-2012, 2012.

Liu, Z., Y. Wang, D. Gu, C. Zhao, L. G. Huey, R. Stickel, J. Liao, M. Shao, T. Zhu, and L. Zeng, Evidence of reactive aromatics as a major source of peroxy acetyl nitrate over China, Environmental science & technology, 44(18), 7017-7022, doi:10.1021/es1007966, 2010.

Liu, Z., Y. Wang, M. Vrekoussis, A. Richter, F. Wittrock, J. P. Burrows, M. Shao, C. C. Chang, S. C. Liu, and H. Wang, Exploring the missing source of glyoxal (CHOCHO) over China, Geophysical Research Letters, 39(10), doi:10.1029/2012GL051645, 2012.

Logan, J. A., M. J. Prather, S. C. Wofsy, and M. B. McElroy, Tropospheric chemistry: A global perspective, Journal of Geophysical Research: Oceans, 86(C8), 7210-7254, doi:10.1029/JC086iC08p07210, 1981.

Luecken, D., W. Hutzell, and G. Gipson, Development and analysis of air quality modeling simulations for hazardous air pollutants, Atmospheric Environment, 40(26), 5087-5096, 2006.

Mao, H., and R. Talbot, O3 and CO in New England: Temporal variations and relationships, Journal of Geophysical Research: Atmospheres, 109(D21), doi:10.1029/2004JD004913, 2004.

Mellqvist, J., J. Samuelsson, J. Johansson, C. Rivera, B. Lefer, S. Alvarez, and J. Jolly, Measurements of industrial emissions of alkenes in Texas using the solar occultation flux method, Journal of Geophysical Research: Atmospheres, 115(D7), 2010.

Milazzo, M. F., G. Ancione, and R. Lisi, Emissions of volatile organic compounds during the ship-loading of petroleum products: Dispersion modelling and environmental concerns, Journal of environmental management, 204, 637-650, 2017.

Millet, D. B., D. J. Jacob, K. F. Boersma, T. M. Fu, T. P. Kurosu, K. Chance, C. L. Heald, and A. Guenther, Spatial distribution of isoprene emissions from North America derived from formaldehyde column measurements by the OMI satellite sensor, Journal of Geophysical Research: Atmospheres, 113(D2), 2008.

Olaguer, E. P., B. Rappenglück, B. Lefer, J. Stutz, J. Dibb, R. Griffin, W. H. Brune, M. Shauck, M. Buhr, and H. Jeffries, Deciphering the role of radical precursors during the Second Texas Air Quality Study, Journal of the Air & Waste Management Association, 59(11), 1258-1277, 2009.

Pang, X., Y. Mu, Y. Zhang, X. Lee, and J. Yuan, Contribution of isoprene to formaldehyde and ozone formation based on its oxidation products measurement in Beijing, China, Atmospheric Environment, 43(13), 2142-2147, doi:10.1016/j.atmosenv.2009.01.022, 2009.

Parrish, D., D. Allen, T. Bates, M. Estes, F. Fehsenfeld, G. Feingold, R. Ferrare, R. Hardesty, J. Meagher, and J. Nielsen-Gammon, Overview of the second Texas air quality study (TexAQS II) and the Gulf of Mexico atmospheric composition and climate study (GoMACCS), Journal of Geophysical Research: Atmospheres, 114(D7), 2009.

Parrish, D., T. Ryerson, J. Mellqvist, J. Johansson, A. Fried, D. Richter, J. Walega, R. d. Washenfelder, J. De Gouw, and J. Peischl, Primary and secondary sources of formaldehyde in urban atmospheres: Houston Texas region, Atmospheric Chemistry and Physics, 12(7), 3273-3288, 2012.

Parrish, D., M. Trainer, D. Hereid, E. Williams, K. Olszyna, R. Harley, J. Meagher, and F. Fehsenfeld, Decadal change in carbon monoxide to nitrogen oxide ratio in US vehicular emissions, Journal of Geophysical Research: Atmospheres, 107(D12), 2002.

Parrish, D., M. Trainer, J. Holloway, J. Yee, M. Warshawsky, F. Fehsenfeld, G. Forbes, and J. Moody, Relationships between ozone and carbon monoxide at surface sites in the North Atlantic region, Journal of Geophysical Research: Atmospheres, 103(D11), 13357-13376, doi:10.1029/98JD00376, 1998.

Parrish, D. D., J. S. Holloway, M. Trainer, P. C. Murphy, F. C. Fehsenfeld, and G. L. Forbes, Export of North American ozone pollution to the north Atlantic Ocean, Science, 259(5100), 1436-1439, doi:10.1126/science.259.5100.1436, 1993.

Perez, P., and J. Reyes, An integrated neural network model for PM10 forecasting, Atmospheric Environment, 40(16), 2845-2851, 2006.

Pétron, G., A. Karion, C. Sweeney, B. R. Miller, S. A. Montzka, G. J. Frost, M. Trainer, P. Tans, A. Andrews, and J. Kofler, A new look at methane and nonmethane hydrocarbon emissions from oil and natural gas operations in the Colorado Denver-Julesburg Basin, Journal of Geophysical Research: Atmospheres, 119(11), 6836-6852, 2014.

Pierce, T., C. Geron, L. Bender, R. Dennis, G. Tonnesen, and A. Guenther, Influence of increased isoprene emissions on regional ozone modeling, Journal of Geophysical Research: Atmospheres, 103(D19), 25611-25629, doi:10.1029/98JD01804, 1998.

Rappenglück, B., P. Dasgupta, M. Leuchner, Q. Li, and W. Luke, Formaldehyde and its relation to CO, PAN, and SO 2 in the Houston-Galveston airshed, Atmospheric Chemistry and Physics, 10(5), 2413-2424, 2010.

Real, E., K. S. Law, H. Schlager, A. Roiger, H. Huntrieser, J. Methven, M. Cain, J. Holloway, J. Neuman, and T. Ryerson, Lagrangian analysis of low altitude anthropogenic plume processing across the North Atlantic, Atmospheric Chemistry and Physics, 8(24), 7737-7754, 2008.

Ryerson, T., M. Trainer, W. Angevine, C. Brock, R. Dissly, F. Fehsenfeld, G. Frost, P. Goldan, J. Holloway, and G. Hübler, Effect of petrochemical industrial emissions of reactive alkenes and NOx on tropospheric ozone formation in Houston, Texas, Journal of Geophysical Research: Atmospheres, 108(D8), 2003.

Sachse, G. W., G. F. Hill, L. O. Wade, and M. G. Perry, Fast-response, high-precision carbon monoxide sensor using a tunable diode laser absorption technique, Journal of Geophysical Research: Atmospheres, 92(D2), 2071-2081, doi:10.1029/JD092iD02p02071, 1987.

Seinfeld, J. H., and S. N. Pandis (2016), Atmospheric chemistry and physics: from air pollution to climate change, John Wiley & Sons.

Shad, R., M. S. Mesgari, and A. Shad, Predicting air pollution using fuzzy genetic linear membership kriging in GIS, Computers, Environment and Urban Systems, 33(6), 472-481, 2009.

Sindelarova, K., C. Granier, I. Bouarar, A. Guenther, S. Tilmes, T. Stavrakou, J.-F. Müller, U. Kuhn, P. Stefani, and W. Knorr, Global data set of biogenic VOC emissions calculated by the MEGAN model over the last 30 years, Atmospheric Chemistry and Physics, 14(17), 9317-9341, 2014.

Song, C., H. Kim, R. von Glasow, P. Brimblecombe, J. Kim, R. Park, and J. Woo, Source identification and budget analysis on elevated levels of formaldehyde within ship plumes: a photochemical/dynamic model analysis, Atmospheric Chemistry & Physics Discussions, 10(6), 2010.

Straume, A., H. Schrijver, A. Gloudemans, S. Houweling, I. Aben, A. Maurellis, A. De Laat, Q. Kleipool, G. Lichtenberg, and R. Van Hees, The global variation of CH4 and CO as seen by SCIAMACHY, Advances in Space Research, 36(5), 821-827, 2005.

Van der Wal, J., and L. Janssen, Analysis of spatial and temporal variations of PM 10 concentrations in the Netherlands using Kalman filtering, Atmospheric Environment, 34(22), 3675-3687, 2000.

Vautard, R., M. Schaap, R. Bergström, B. Bessagnet, J. Brandt, P. Builtjes, J. Christensen, C. Cuvelier, V. Foltescu, and A. Graff, Skill and uncertainty of a regional air quality model ensemble, Atmospheric Environment, 43(31), 4822-4832, 2009.

Volkamer, R., L. T. Molina, M. J. Molina, T. Shirley, and W. H. Brune, DOAS measurement of glyoxal as an indicator for fast VOC chemistry in urban air, Geophysical Research Letters, 32(8), 2005.

Voulgarakis, A., P. Telford, A. Aghedo, P. Braesicke, G. Faluvegi, N. Abraham, K. Bowman, J. Pyle, and D. Shindell, Global multi-year O 3-CO correlation patterns from models and TES satellite observations, Atmospheric Chemistry and Physics, 11(12), 5819-5838, 2011.

Wang, Y., Y. Choi, T. Zeng, D. Davis, M. Buhr, L. G. Huey, and W. Neff, Assessing the photochemical impact of snow NOx emissions over Antarctica during ANTCI 2003, Atmospheric Environment, 41(19), 3944-3958, doi:10.1016/j.atmosenv.2007.01.056, 2007.

Wang, Y., Y. Choi, T. Zeng, B. Ridley, N. Blake, D. Blake, and F. Flocke, Late-spring increase of trans-Pacific pollution transport in the upper troposphere, Geophysical Research Letters, 33(1), doi:10.1029/2005GL024975, 2006.

Wang, Y., and D. J. Jacob, Anthropogenic forcing on tropospheric ozone and OH since preindustrial times, doi:10.1029/1998JD100004, 1998.

Wang, Y., D. J. Jacob, and J. A. Logan, Global simulation of tropospheric O3-NO x-hydrocarbon chemistry: 1. Model formulation, Journal of Geophysical Research: Atmospheres, 103(D9), 10713-10725, doi:10.1029/98JD00158, 1998.

Wang, Y., and T. Zeng, On tracer correlations in the troposphere: The case of ethane and propane, Journal of Geophysical Research: Atmospheres, 109(D24), doi:10.1029/2004JD005023, 2004.

Washenfelder, R., M. Trainer, G. Frost, T. Ryerson, E. Atlas, J. De Gouw, F. Flocke, A. Fried, J. Holloway, and D. Parrish, Characterization of NOx, SO2, ethene, and propene from industrial emission sources in Houston, Texas, Journal of Geophysical Research: Atmospheres, 115(D16), 2010.

Weibring, P., D. Richter, J. Walega, L. Rippe, and A. Fried, Difference frequency generation spectrometer for simultaneous multispecies detection, Optics express, 18(26), 27670-27681, doi:10.1364/OE.18.027670, 2010.

Wert, B., M. Trainer, A. Fried, T. Ryerson, B. Henry, W. Potter, W. Angevine, E. Atlas, S. Donnelly, and F. Fehsenfeld, Signatures of terminal alkene oxidation in airborne

formaldehyde measurements during TexAQS 2000, Journal of Geophysical Research: Atmospheres, 108(D3), 2003.

Wiedinmyer, C., S. Friedfeld, W. Baugh, J. Greenberg, A. Guenther, M. Fraser, and D. Allen, Measurement and analysis of atmospheric concentrations of isoprene and its reaction products in central Texas, Atmospheric Environment, 35(6), 1001-1013, 2001.

Wittrock, F., A. Richter, H. Oetjen, J. P. Burrows, M. Kanakidou, S. Myriokefalitakis, R. Volkamer, S. Beirle, U. Platt, and T. Wagner, Simultaneous global observations of glyoxal and formaldehyde from space, Geophysical Research Letters, 33(16), 2006.

Xie, Y., and C. M. Berkowitz, The use of positive matrix factorization with conditional probability functions in air quality studies: an application to hydrocarbon emissions in Houston, Texas, Atmospheric Environment, 40(17), 3070-3091, 2006.

Yang, Q., Y. Wang, C. Zhao, Z. Liu, W. I. Gustafson Jr, and M. Shao, NO x emission reduction and its effects on ozone during the 2008 Olympic Games, Environmental science & technology, 45(15), 6404-6410, doi:10.1021/es200675v, 2011.

Zeng, T., and Y. Wang, Nationwide summer peaks of OC/EC ratios in the contiguous United States, Atmospheric environment, 45(3), 578-586, 2011.

Zeng, T., Y. Wang, K. Chance, N. Blake, D. Blake, and B. Ridley, Halogen-driven low-altitude O3 and hydrocarbon losses in spring at northern high latitudes, Journal of Geophysical Research: Atmospheres, 111(D17), doi:10.1029/2005JD006706, 2006.

Zeng, T., Y. Wang, K. Chance, E. V. Browell, B. A. Ridley, and E. L. Atlas, Widespread persistent near-surface ozone depletion at northern high latitudes in spring, Geophysical research letters, 30(24), doi:10.1029/2003GL018587, 2003.

Zhang, L., D. J. Jacob, K. W. Bowman, J. A. Logan, S. Turquety, R. C. Hudman, Q. Li, R. Beer, H. M. Worden, and J. R. Worden, Ozone-CO correlations determined by the TES satellite instrument in continental outflow regions, Geophysical Research Letters, 33(18), 2006.

Zhang, L., H. Jiang, X. Lu, and J. Jin, Comparison analysis of global carbon monoxide concentration derived from SCIAMACHY, AIRS, and MOPITT, International journal of remote sensing, 37(21), 5155-5175, 2016.

Zhang, R., Y. Wang, Q. He, L. Chen, Y. Zhang, H. Qu, C. Smeltzer, J. Li, L. Alvarado, and M. Vrekoussis, Enhanced trans-Himalaya pollution transport to the Tibetan Plateau by cut-off low systems, Atmospheric Chemistry and Physics, 17(4), 3083-3095, 2017.

Zhang, Y., M. Bocquet, V. Mallet, C. Seigneur, and A. Baklanov, Real-time air quality forecasting, part I: History, techniques, and current status, Atmospheric Environment, 60, 632-655, 2012.

Zhang, Y., M. Bocquet, V. Mallet, C. Seigneur, and A. Baklanov, Real-time air quality forecasting, part II: State of the science, current research needs, and future prospects, Atmospheric Environment, 60, 656-676, 2012.

Zhang, Y., and Y. Wang, Climate-driven ground-level ozone extreme in the fall over the Southeast United States, Proceedings of the National Academy of Sciences, 201602563, doi:10.1073/pnas.1602563113, 2016.

Zhang, Y., Y. Wang, G. Chen, C. Smeltzer, J. Crawford, J. Olson, J. Szykman, A. J. Weinheimer, D. J. Knapp, and D. D. Montzka, Large vertical gradient of reactive nitrogen oxides in the boundary layer: Modeling analysis of DISCOVER-AQ 2011 observations, Journal of Geophysical Research: Atmospheres, doi:10.1002/2015JD024203, 2016.

Zhao, C., and Y. Wang, Assimilated inversion of NOx emissions over east Asia using OMI NO2 column measurements, Geophysical Research Letters, 36(6), doi:10.1029/2008GL037123, 2009.

Zhao, C., Y. Wang, Y. Choi, and T. Zeng, Summertime impact of convective transport and lightning NO x production over North America: modeling dependence on meteorological simulations, Atmospheric Chemistry and Physics, 9(13), 4315-4327, doi:10.5194/acp-9-4315-2009, 2009.

Zhao, C., Y. Wang, Q. Yang, R. Fu, D. Cunnold, and Y. Choi, Impact of East Asian summer monsoon on the air quality over China: View from space, Journal of Geophysical Research: Atmospheres, 115(D9), doi:10.1029/2009JD012745, 2010.

Zhao, C., Y. Wang, and T. Zeng, East China plains: A “basin” of ozone pollution, Environmental science & technology, 43(6), 1911-1915, doi:10.1021/es8027764, 2009.

Zhu, L., D. J. Jacob, L. J. Mickley, E. A. Marais, D. S. Cohan, Y. Yoshida, B. N. Duncan, G. G. Abad, and K. V. Chance, Anthropogenic emissions of highly reactive volatile organic compounds in eastern Texas inferred from oversampling of satellite (OMI) measurements of HCHO columns, Environmental Research Letters, 9(11), 114004, 2014.